



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>7</sup> : <b>C12N 15/12, C07K 14/47, C12N 9/12, 5/10, C07K 16/18, A61K 38/17</b></p>	<b>A2</b>	<p>(11) International Publication Number: <b>WO 00/06728</b></p> <p>(43) International Publication Date: <b>10 February 2000 (10.02.00)</b></p>																																																																		
<p>(21) International Application Number: <b>PCT/US99/17132</b></p> <p>(22) International Filing Date: <b>28 July 1999 (28.07.99)</b></p>																																																																				
<p>(30) Priority Data:</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 30%;">09/123,494</td> <td style="width: 40%;">28 July 1998 (28.07.98)</td> <td style="width: 30%;">US</td> </tr> <tr> <td>Not furnished</td> <td>28 July 1998 (28.07.98)</td> <td>US</td> </tr> <tr> <td>09/152,814</td> <td>14 September 1998 (14.09.98)</td> <td>US</td> </tr> <tr> <td>Not furnished</td> <td>14 September 1998 (14.09.98)</td> <td>US</td> </tr> <tr> <td>09/173,482</td> <td>14 October 1998 (14.10.98)</td> <td>US</td> </tr> <tr> <td>Not furnished</td> <td>14 October 1998 (14.10.98)</td> <td>US</td> </tr> <tr> <td>60/106,889</td> <td>3 November 1998 (03.11.98)</td> <td>US</td> </tr> <tr> <td>60/109,093</td> <td>19 November 1998 (19.11.98)</td> <td>US</td> </tr> <tr> <td>60/113,796</td> <td>22 December 1998 (22.12.98)</td> <td>US</td> </tr> <tr> <td>09/229,005</td> <td>12 January 1999 (12.01.99)</td> <td>US</td> </tr> <tr> <td>Not furnished</td> <td>12 January 1999 (12.01.99)</td> <td>US</td> </tr> </table>	09/123,494	28 July 1998 (28.07.98)	US	Not furnished	28 July 1998 (28.07.98)	US	09/152,814	14 September 1998 (14.09.98)	US	Not furnished	14 September 1998 (14.09.98)	US	09/173,482	14 October 1998 (14.10.98)	US	Not furnished	14 October 1998 (14.10.98)	US	60/106,889	3 November 1998 (03.11.98)	US	60/109,093	19 November 1998 (19.11.98)	US	60/113,796	22 December 1998 (22.12.98)	US	09/229,005	12 January 1999 (12.01.99)	US	Not furnished	12 January 1999 (12.01.99)	US	<p>(72) Inventors; and (75) Inventors/Applicants (for US only): <b>HILLMAN, Jennifer, L.</b> [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). <b>LAL, Preeti</b> [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). <b>TANG, Y., Tom</b> [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). <b>CORLEY, Neil, C.</b> [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). <b>GUEGLER, Karl, J.</b> [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). <b>BAUGHN, Mariah, R.</b> [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). <b>PATTERSON, Chandra</b> [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). <b>BANDMAN, Olga</b> [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). <b>AU-YOUNG, Janice</b> [US/US]; 1419 Kains Avenue, Berkeley, CA 94709 (US). <b>GORGONE, Gina, A.</b> [US/US]; 1253 Pinecrest Drive, Boulder Creek, CA 95006 (US). <b>YUE, Henry</b> [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). <b>AZIMZAI, Yalda</b> [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). <b>REDDY, Roopa</b> [IN/US]; 1233 W. McKinley Drive, Sunnyvale, CA 94086 (US). <b>LU, Dyung, Aina, M.</b> [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). <b>SHIH, Leo, L.</b> [US/US]; Apt. B, 1081 Tanland Drive, Palo Alto, CA 94303 (US).</p>																																		
09/123,494	28 July 1998 (28.07.98)	US																																																																		
Not furnished	28 July 1998 (28.07.98)	US																																																																		
09/152,814	14 September 1998 (14.09.98)	US																																																																		
Not furnished	14 September 1998 (14.09.98)	US																																																																		
09/173,482	14 October 1998 (14.10.98)	US																																																																		
Not furnished	14 October 1998 (14.10.98)	US																																																																		
60/106,889	3 November 1998 (03.11.98)	US																																																																		
60/109,093	19 November 1998 (19.11.98)	US																																																																		
60/113,796	22 December 1998 (22.12.98)	US																																																																		
09/229,005	12 January 1999 (12.01.99)	US																																																																		
Not furnished	12 January 1999 (12.01.99)	US																																																																		
<p>(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 30%;">US</td> <td style="width: 40%;">Not furnished (CIP)</td> <td style="width: 30%;"></td> </tr> <tr> <td>Filed on</td> <td>28 July 1998 (28.07.98)</td> <td></td> </tr> <tr> <td>US</td> <td>09/123,494 (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>28 July 1998 (28.07.98)</td> <td></td> </tr> <tr> <td>US</td> <td>09/152,814 (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>14 September 1998 (14.09.98)</td> <td></td> </tr> <tr> <td>US</td> <td>Not furnished (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>14 September 1998 (14.09.98)</td> <td></td> </tr> <tr> <td>US</td> <td>09/173,482 (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>14 October 1998 (14.10.98)</td> <td></td> </tr> <tr> <td>US</td> <td>Not furnished (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>14 October 1998 (14.10.98)</td> <td></td> </tr> <tr> <td>US</td> <td>60/106,889 (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>3 November 1998 (03.11.98)</td> <td></td> </tr> <tr> <td>US</td> <td>60/109,093 (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>19 November 1998 (19.11.98)</td> <td></td> </tr> <tr> <td>US</td> <td>60/113,796 (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>22 December 1998 (22.12.98)</td> <td></td> </tr> <tr> <td>US</td> <td>09/229,005 (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>12 January 1999 (12.01.99)</td> <td></td> </tr> <tr> <td>US</td> <td>Not furnished (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>12 January 1999 (12.01.99)</td> <td></td> </tr> </table>	US	Not furnished (CIP)		Filed on	28 July 1998 (28.07.98)		US	09/123,494 (CIP)		Filed on	28 July 1998 (28.07.98)		US	09/152,814 (CIP)		Filed on	14 September 1998 (14.09.98)		US	Not furnished (CIP)		Filed on	14 September 1998 (14.09.98)		US	09/173,482 (CIP)		Filed on	14 October 1998 (14.10.98)		US	Not furnished (CIP)		Filed on	14 October 1998 (14.10.98)		US	60/106,889 (CIP)		Filed on	3 November 1998 (03.11.98)		US	60/109,093 (CIP)		Filed on	19 November 1998 (19.11.98)		US	60/113,796 (CIP)		Filed on	22 December 1998 (22.12.98)		US	09/229,005 (CIP)		Filed on	12 January 1999 (12.01.99)		US	Not furnished (CIP)		Filed on	12 January 1999 (12.01.99)		<p>(74) Agents: <b>BILLINGS, Lucy, J. et al.</b>; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).</p> <p>(81) Designated States: <b>AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</b></p>	
US	Not furnished (CIP)																																																																			
Filed on	28 July 1998 (28.07.98)																																																																			
US	09/123,494 (CIP)																																																																			
Filed on	28 July 1998 (28.07.98)																																																																			
US	09/152,814 (CIP)																																																																			
Filed on	14 September 1998 (14.09.98)																																																																			
US	Not furnished (CIP)																																																																			
Filed on	14 September 1998 (14.09.98)																																																																			
US	09/173,482 (CIP)																																																																			
Filed on	14 October 1998 (14.10.98)																																																																			
US	Not furnished (CIP)																																																																			
Filed on	14 October 1998 (14.10.98)																																																																			
US	60/106,889 (CIP)																																																																			
Filed on	3 November 1998 (03.11.98)																																																																			
US	60/109,093 (CIP)																																																																			
Filed on	19 November 1998 (19.11.98)																																																																			
US	60/113,796 (CIP)																																																																			
Filed on	22 December 1998 (22.12.98)																																																																			
US	09/229,005 (CIP)																																																																			
Filed on	12 January 1999 (12.01.99)																																																																			
US	Not furnished (CIP)																																																																			
Filed on	12 January 1999 (12.01.99)																																																																			
<p>(71) Applicant (for all designated States except US): <b>INCYTE PHARMACEUTICALS, INC.</b> [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).</p>																																																																				
<p>(54) Title: <b>PHOSPHORYLATION EFFECTORS</b></p>																																																																				
<p>(57) Abstract</p> <p>The invention provides human phosphorylation effectors (PHSP) and polynucleotides which identify and encode PHSP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of PHSP.</p>																																																																				

**Published**

*Without international search report and to be republished upon receipt of that report.*

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## PHOSPHORYLATION EFFECTORS

### TECHNICAL FIELD

5           This invention relates to nucleic acid and amino acid sequences of phosphorylation effectors and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, immune, and neuronal disorders.

          Kinases and phosphatases are critical components of intracellular signal transduction mechanisms. Kinases catalyze the transfer of high energy phosphate groups from adenosine triphosphate (ATP) to various target proteins. Phosphatases, in contrast, remove phosphate groups from proteins. Reversible protein phosphorylation is the main strategy for regulating protein activity in eukaryotic cells. In general, proteins are activated by phosphorylation in response to extracellular signals such as hormones, neurotransmitters, and growth and differentiation factors.

15   Protein dephosphorylation occurs when down-regulation of a signaling pathway is required. The coordinate activities of kinases and phosphatases regulate key cellular processes such as proliferation, differentiation, and cell cycle progression. Kinases comprise the largest known enzyme superfamily and are widely varied in their substrate specificities. Kinases may be categorized based on the specific amino acid residues that are phosphorylated in their substrates:

20   protein tyrosine kinases (PTK) phosphorylate tyrosine residues, and protein serine/threonine kinases (STK) phosphorylate serine and/or threonine residues. Almost all kinases contain a conserved 250-300 amino acid catalytic domain. This domain can be further divided into 11 subdomains. N-terminal subdomains I-IV fold into a two-lobed structure which binds and orients the ATP donor molecule, and subdomain V spans the two lobes. C-terminal subdomains VIA-XI

25   bind the protein substrate and transfer the gamma phosphate from ATP to the hydroxyl group of a serine, threonine, or tyrosine residue. Each of the 11 subdomains contains specific catalytic residues or amino acid motifs characteristic of that subdomain. For example, subdomain I contains an 8-amino acid glycine-rich ATP binding consensus motif, subdomain II contains a critical lysine residue required for maximal catalytic activity, and subdomains VI and IX comprise

30   the highly conserved catalytic core. Kinases may also be categorized by additional amino acid sequences, generally between 5 and 100 residues, which either flank or occur within the kinase domain. These additional amino acid sequences regulate kinase activity and determine substrate specificity. (Reviewed in Hardie, G. and Hanks, S. (1995) The Protein Kinase Facts Books, Vol I:7-20 Academic Press, San Diego, CA.)

35           STKs include both protein kinase A (PKA) and calcium-dependent protein kinase C

(PKC), both of which transduce signals from plasma membrane receptors. The activities of PKA and PKC are directly regulated by second messenger signaling molecules such as cyclic AMP and diacylglycerol, respectively. A novel kinase identified by genetic analysis in the fission yeast Schizosaccharomyces pombe is encoded by the *cek1*<sup>+</sup> gene and is related to both PKA and PKC

5 (Samejima, I. and Yanagida, M. (1994) Mol. Cell. Biol. 14:6361-6371). *cek1*<sup>+</sup> encodes an unusually large kinase of 1309 amino acids. The kinase domain spans residues 585 to 987, and 112 additional amino acids are present in this domain between subdomains VII and VIII. Overexpression of *cek1*<sup>+</sup> suppresses mutations in *cut8*<sup>+</sup>, a gene required for chromosome segregation during mitosis. Therefore, *cek1*<sup>+</sup> may encode a unique member of the PKA/PKC

10 protein family with a role in mitotic signaling and cell cycle progression.

PTKs may be classified as either transmembrane or nontransmembrane proteins.

Transmembrane tyrosine kinases function as receptors for most growth factors. Binding of growth factor to the receptor activates the transfer of a phosphate group from ATP to selected tyrosine side chains of the receptor itself and other specific second messenger proteins. Growth factors

15 (GF) that associate with receptor PTKs include epidermal GF, platelet-derived GF, fibroblast GF, hepatocyte GF, insulin and insulin-like GFs, nerve GF, vascular endothelial GF, and macrophage colony stimulating factor. Nontransmembrane PTKs form signaling complexes with the cytosolic domains of plasma membrane receptors. Receptors that signal through nontransmembrane PTKs include cytokine, hormone, and antigen-specific lymphocytic receptors. Many PTKs were first

20 identified as oncogene products in cancer cells in which PTK activation was no longer subject to normal cellular controls. In fact, about one third of the known oncogenes encode PTKs. Furthermore, cellular transformation (oncogenesis) is often accompanied by increased tyrosine phosphorylation activity (Charbonneau, H. and Tonks, N. K. (1992) Annu. Rev. Cell Biol. 8:463-93). Regulation of PTK activity may therefore be an important strategy in controlling some

25 types of cancer.

Some kinases utilize carbohydrates as their substrates and are important for glucose metabolism. For example, glycolysis employs four distinct kinases to effect the conversion of glucose to pyruvate, a key metabolite in the production of ATP. One of these enzymes is phosphofructokinase (PFK) which catalyzes the transfer of phosphate from ATP to fructose 6-

30 phosphate. PFK is an allosteric enzyme and a key regulator of glycolysis. In certain genetic muscle disorders, such as muscle phosphofructokinase deficiency type VII, phosphofructokinase activity is absent in muscle and deficient in red blood cells. As a result, afflicted individuals suffer from mild hemolytic anemia and muscle pain (Isselbacher, K.J. et al. (1994) Harrison's Principles of Internal Medicine, McGraw-Hill, New York, NY, p. 2102).

35 Kinase-mediated phosphorylation is antagonized by the activity of phosphatases, which

remove phosphate groups by hydrolysis. Phosphatases are classified into one of three evolutionarily distinct families: the protein serine/threonine phosphatases (PPs), the protein tyrosine phosphatases, and the acid/alkaline phosphatases. PPs may be further categorized into four distinct groups: PP-I, PP-IIA, PP-IIB, and PP-IIC. (Cohen, P. (1989) Annu. Rev. Biochem. 58:453-508). PP-I, in particular, dephosphorylates many of the proteins phosphorylated by PKA and is therefore an important regulator of signal transduction pathways. Kinase-activated proteins which bind to and inhibit PP-I have been identified. These inhibitors potentiate the activity of kinases such as PKA by allowing protein substrates to remain in their phosphorylated, activated state. A novel inhibitor of PP-1 has been purified from porcine aorta (Eto, M. et al. (1995) J. Biochem. 118:1104-1107; Eto, M. et al. (1997) FEBS Lett. 410:356-360). This inhibitor, called CPI17, is 147 amino acids in length and is activated by PKC. CPI17 expression is restricted to smooth muscle tissues such as aorta and bladder, suggesting that CPI17 functions in PKC-mediated signal transduction pathways in these tissues, possibly through a calcium-dependent mechanism.

15       The discovery of new phosphorylation effectors and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, immune, and neuronal disorders.

#### SUMMARY OF THE INVENTION

20

The invention features substantially purified polypeptides, phosphorylation effectors, referred to collectively as "PHSP" and individually as "PHSP-1 to PHSP-31". In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof.

25       The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-31, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof. The invention also includes an

30       isolated and purified polynucleotide variant having at least 80% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof.

          Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising

35       an amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments

thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample  
5 containing nucleic acids, the method comprising the steps of (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

10 The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:32-62, and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 80% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:32-62, and fragments thereof. The invention also provides an  
15 isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:32-62, and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the  
20 group consisting of SEQ ID NO:1-31, and fragments thereof. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and  
25 (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected  
30 from the group consisting of SEQ ID NO:1-31, and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of PHSP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a  
35 substantially purified polypeptide having the amino acid sequence selected from the group

consisting of SEQ ID NO:1-31, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of PHSP, the method comprising administering to a subject in  
5 need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof.

### BRIEF DESCRIPTION OF THE TABLES

Table 1 shows nucleotide and polypeptide sequence identification numbers (SEQ ID NO),  
10 clone identification numbers (clone ID), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding PHSP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods and algorithms used for identification of PHSP.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence as  
15 determined by northern analysis, diseases, disorders, or conditions associated with these tissues, and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding PHSP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze PHSP, along with  
20 applicable descriptions, references, and threshold parameters.

### DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods  
25 described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for  
30 example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention  
35 belongs. Although any machines, materials, and methods similar or equivalent to those described

herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

## DEFINITIONS

"PHSP" refers to the amino acid sequences of substantially purified PHSP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which, when bound to PHSP, increases or prolongs the duration of the effect of PHSP. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of PHSP.

An "allelic variant" is an alternative form of the gene encoding PHSP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding PHSP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as PHSP or a polypeptide with at least one functional characteristic of PHSP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding PHSP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding PHSP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent PHSP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of PHSP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine,



and valine; glycine and alanine; asparagine and glutamin ; serine and threonine; and phenylalanine and tyrosine.

The terms "amino acid" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of PHSP which are preferably at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of PHSP. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which, when bound to PHSP, decreases the amount or the duration of the effect of the biological or immunological activity of PHSP. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of PHSP.

The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, F(ab')<sub>2</sub>, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind PHSP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell,

the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active," refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic PHSP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" or a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding PHSP or fragments of PHSP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

The term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding PHSP, by northern analysis is indicative of the presence of nucleic acids encoding PHSP in a sample, and

thereby correlates with expression of the transcript from the polynucleotide encoding PHSP.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a  
5 polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide  
10 from which it was derived.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of  
15 hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced  
20 stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not  
25 hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Madison WI) which creates alignments between two or more sequences according to methods  
30 selected by the user, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the  
35 number of gap residues in sequence B, into the sum of the residue matches between sequence A

and sequence B, times on hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between  
5 sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

“Human artificial chromosomes” (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

10 The term “humanized antibody” refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

“Hybridization” refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

15 The term “hybridization complex” refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C<sub>0</sub>t or R<sub>0</sub>t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate  
20 substrate to which cells or their nucleic acids have been fixed).

The words “insertion” or “addition” refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

“Immune response” can refer to conditions associated with inflammation, trauma, immune  
25 disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term “microarray” refers to an arrangement of distinct polynucleotides on a substrate.

The terms “element” or “array element” in a microarray context, refer to hybridizable  
30 polynucleotides arranged on the surface of a substrate.

The term “modulate” refers to a change in the activity of PHSP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of PHSP.

The phrases “nucleic acid” or “nucleic acid sequence,” as used herein, refer to a  
35 nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to

DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, "fragments" refers to those nucleic acid sequences which, comprise a region of unique polynucleotide sequence that specifically identifies  
5 SEQ ID NO:32-62, for example, as distinct from any other sequence in the same genome. For example, a fragment of SEQ ID NO:32-62 is useful in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:32-62 from related polynucleotide sequences. A fragment of SEQ ID NO:32-62 is at least about 15-20 nucleotides in length. The precise length of the fragment of SEQ ID NO:32-62 and the region of SEQ ID  
10 NO:32-62 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment. In some cases, a fragment, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length polypeptide

The terms "operably associated" or "operably linked" refer to functionally related nucleic  
15 acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

20 The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

25 "Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

30 The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding PHSP, or fragments thereof, or PHSP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" or "specifically binding" refer to that interaction between a  
35 protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon

the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent, e.g., formamide, temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of PHSP polypeptides refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g., replacement of

glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

- 5           The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to PHSP. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The
- 10           corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide
- 15           polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

#### **THE INVENTION**

- The invention is based on the discovery of new human phosphorylation effectors (PHSP),
- 20           the polynucleotides encoding PHSP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, immune, and neuronal disorders.

- Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding PHSP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte
- 25           clones in which nucleic acids encoding each PHSP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The clones in column 5 were used to assemble the consensus nucleotide sequence of each PHSP and are useful as fragments in hybridization technologies.

- 30           The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO and column 2 shows the number of amino acid residues in each polypeptide. Columns 3 and 4 show potential phosphorylation sites and potential glycosylation sites, respectively. Column 5 shows the amino acid residues comprising signature sequences and motifs. Column 6 shows homologous sequences as identified by BLAST analysis,
- 35           while column 7 shows analytical methods used to identify each polypeptide through sequence

homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding PHSP. The first column of Table 3 lists the SEQ ID NOs. Column 2 lists tissue categories which express PHSP as a fraction of total tissue categories expressing PHSP. Column 3 lists diseases, disorders, or conditions associated with those tissues expressing PHSP. Column 4 lists the vectors used to subclone the cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding PHSP were isolated. Column 1 references the SEQ ID NO, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The following fragments of the nucleotide sequences encoding PHSP are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:32-62 and to distinguish between SEQ ID NO:32-62 and related polynucleotide sequences. The useful fragments include, the fragment of SEQ ID NO:32 from about nucleotide 81 to about nucleotide 110; the fragment of SEQ ID NO:33 from about nucleotide 323 to about nucleotide 352; the fragment of SEQ ID NO:34 from about nucleotide 83 to about nucleotide 112; the fragment of SEQ ID NO:35 from about nucleotide 524 to about nucleotide 553; the fragment of SEQ ID NO:36 from about nucleotide 275 to about nucleotide 346; the fragment of SEQ ID NO:37 from about nucleotide 1328 to about nucleotide 1396; the fragment of SEQ ID NO:38 from about nucleotide 245 to about nucleotide 304; the fragment of SEQ ID NO:39 from about nucleotide 1253 to about nucleotide 1312; the fragment of SEQ ID NO:41 from about nucleotide 117 to about nucleotide 170; the fragments of SEQ ID NO:42 from about nucleotide 109 to about nucleotide 153, and from about nucleotide 325 to about nucleotide 369; the fragments of SEQ ID NO:43 from about nucleotide 380 to about nucleotide 424, and from about nucleotide 1190 to about nucleotide 1234; the fragment of SEQ ID NO:44 from about nucleotide 1 to about nucleotide 46; the fragment of SEQ ID NO:45 from about nucleotide 533 to about nucleotide 577; the fragments of SEQ ID NO:46 from about nucleotide 109 to about nucleotide 153, and from about nucleotide 379 to about nucleotide 423; the fragment of SEQ ID NO:47 from about nucleotide 1730 to about nucleotide 1774; the fragment of SEQ ID NO:48 from about nucleotide 433 to about nucleotide 477; the fragment of SEQ ID NO:49 from about nucleotide 1117 to about nucleotide 1155; the fragment of SEQ ID NO:50 from about nucleotide 166 to about nucleotide 213; the fragment of SEQ ID NO:51 from about nucleotide 60 to about nucleotide 95; the fragment of SEQ ID NO:52 from about nucleotide 326 to about nucleotide 370; the fragment of SEQ ID NO:53 from about nucleotide 25 to about nucleotide 66; the fragment of SEQ ID NO:54 from about nucleotide 55 to



about nucleotide 102; the fragment of SEQ ID NO:55 from about nucleotide 138 to about nucleotide 167; the fragment of SEQ ID NO:56 from about nucleotide 29 to about nucleotide 58; the fragment of SEQ ID NO:57 from about nucleotide 455 to about nucleotide 484; the fragment of SEQ ID NO:58 from about nucleotide 226 to about nucleotide 255; the fragment of SEQ ID NO:59 from about nucleotide 557 to about nucleotide 598; the fragment of SEQ ID NO:60 from about nucleotide 284 to about nucleotide 325; the fragment of SEQ ID NO:61 from about nucleotide 1043 to about nucleotide 1090; and the fragment of SEQ ID NO:62 from about nucleotide 84 to about nucleotide 132. The polypeptides encoded by the fragments of SEQ ID NO:32-62 are useful, for example, as immunogenic peptides.

10       The invention also encompasses PHSP variants. A preferred PHSP variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the PHSP amino acid sequence, and which contains at least one functional or structural characteristic of PHSP.

15       The invention also encompasses polynucleotides which encode PHSP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:32-62, which encodes PHSP.

20       The invention also encompasses a variant of a polynucleotide sequence encoding PHSP. In particular, such a variant polynucleotide sequence will have at least about 80%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding PHSP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:32-62 which has at least about 80%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:32-62. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of PHSP.

25       It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding PHSP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring PHSP, and all such variations are to be considered as being specifically disclosed.

30       Although nucleotide sequences which encode PHSP and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring PHSP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding

PHSP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide  
5 sequence encoding PHSP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode PHSP and PHSP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the  
10 synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding PHSP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID  
15 NO:32-62 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low  
20 stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the  
25 concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35%  
30 formamide, and 100 µg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 µg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency  
35 conditions can be defined by salt concentration and by temperature. As above, wash stringency can

be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least  
5 about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will  
10 be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations  
15 of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 (Hamilton, Reno NV), Peltier thermal cycler 200 (PTC200; MJ Research, Watertown MA) and the ABI CATALYST 800 (Perkin-Elmer). Sequencing is then carried out using the ABI 373 or 377 DNA sequencing systems (Perkin-Elmer), or the MEGABACE  
20 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

25 The nucleic acid sequences encoding PHSP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.)  
30 Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, .g., Lagerstrom, M. et  
35 al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions

and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-306). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode PHSP may be cloned in recombinant DNA molecules that direct expression of PHSP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express PHSP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter PHSP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction

sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding PHSP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, and Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225-232.)

5 Alternatively, PHSP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of PHSP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other  
10 proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New  
15 York NY.)

In order to express a biologically active PHSP, the nucleotide sequences encoding PHSP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and  
20 inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding PHSP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding PHSP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding PHSP and its initiation codon and upstream regulatory sequences are inserted into  
25 the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion  
30 of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding PHSP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and  
35 in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory

Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding PHSP. These include, but are not limited to, microorganisms such as bacteria transformed  
5 with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

10 In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding PHSP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding PHSP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or pSPORT1 plasmid (Life Technologies). Ligation of sequences encoding PHSP into the vector's multiple cloning site  
15 disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509.) When large quantities of PHSP are needed, e.g. for the production of antibodies,  
20 vectors which direct high level expression of PHSP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of PHSP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct  
25 either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Grant et al. (1987) *Methods Enzymol.* 153:516-54; and Scorer, C. A. et al. (1994) *Bio/Technology* 12:181-184.)

Plant systems may also be used for expression of PHSP. Transcription of sequences encoding PHSP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in  
30 combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See,  
35 e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY,

pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding PHSP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader  
5 sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses PHSP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

10 Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355.)

15 For long term production of recombinant proteins in mammalian systems, stable expression of PHSP in cell lines is preferred. For example, sequences encoding PHSP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before  
20 being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These  
25 include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* or *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides, neomycin and G-418; and *als* or *pat*  
30 confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP;  
35 Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may

be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) *Methods Mol. Biol.* 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding PHSP is inserted within a marker gene sequence, transformed cells containing sequences encoding PHSP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding PHSP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding PHSP and that express PHSP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of PHSP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on PHSP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul MN, Sect. IV; Coligan, J. E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding PHSP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding PHSP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for



ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding PHSP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein  
5 produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode PHSP may be designed to contain signal sequences which direct secretion of PHSP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the  
10 inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational  
15 activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Bethesda MD) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding PHSP may be ligated to a heterologous sequence resulting in translation of a  
20 fusion protein in any of the aforementioned host systems. For example, a chimeric PHSP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of PHSP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST),  
25 maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies  
30 that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the PHSP encoding sequence and the heterologous protein sequence, so that PHSP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch 10). A variety of commercially available kits may also be used to facilitate expression and purification of  
35 fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled PHSP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract systems (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably <sup>35</sup>S-methionine.

Fragments of PHSP may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Various fragments of PHSP may be synthesized separately and then combined to produce the full length molecule.

### THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of PHSP and protein phosphatases. In addition, the expression of PHSP is closely associated with reproductive tissue, nervous tissue, gastrointestinal tissue, cell proliferation, cancer, inflammation, and immune response. Therefore, PHSP appears to play a role in cell proliferative, immune, and neuronal disorders. In the treatment of disorders associated with increased PHSP expression or activity, it is desirable to decrease the expression or activity of PHSP. In the treatment of disorders associated with decreased PHSP expression or activity, it is desirable to increase the expression or activity of PHSP.

Therefore, in one embodiment, PHSP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PHSP. Examples of such disorders include, but are not limited to, a cell proliferative disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis,

hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner  
5 syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a neuronal disorder, such as akathisia, Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, bipolar disorder, catatonia, dementia, depression, diabetic neuropathy, Down's syndrome, tardive dyskinesia, dystonias, epilepsy, Huntington's disease, peripheral neuropathy, multiple sclerosis, neurofibromatosis, Parkinson's  
10 disease, paranoid psychoses, postherpetic neuralgia, schizophrenia, and Tourette's disorder.

In another embodiment, a vector capable of expressing PHSP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PHSP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified  
15 PHSP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PHSP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of PHSP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity  
20 of PHSP including, but not limited to, those listed above.

In a further embodiment, an antagonist of PHSP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of PHSP. Examples of such disorders include, but are not limited to, those described above. In one aspect, an antibody which specifically binds PHSP may be used directly as an antagonist or indirectly as a targeting or delivery  
25 mechanism for bringing a pharmaceutical agent to cells or tissue which express PHSP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding PHSP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of PHSP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary  
30 sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic  
35 efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of PHSP may be produced using methods which are generally known in the art. In particular, purified PHSP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind PHSP. Antibodies to PHSP may also be generated using methods that are well known in the art. Such antibodies may include, but are not  
5 limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with PHSP or with any fragment or oligopeptide thereof  
10 which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

15 It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to PHSP have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of PHSP amino acids may be  
20 fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to PHSP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma  
25 technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate  
30 antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce PHSP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be  
35 generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton

D.R. (1991) Proc. Natl. Acad. Sci. 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; 5 Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for PHSP may also be generated. For example, such fragments include, but are not limited to, F(ab')<sub>2</sub> fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and 10 easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such 15 immunoassays typically involve the measurement of complex formation between PHSP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering PHSP epitopes is preferred, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay 20 techniques may be used to assess the affinity of antibodies for PHSP. Affinity is expressed as an association constant,  $K_a$ , which is defined as the molar concentration of PHSP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The  $K_a$  determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple PHSP epitopes, represents the average affinity, or avidity, of the antibodies for 25 PHSP. The  $K_a$  determined for a preparation of monoclonal antibodies, which are monospecific for a particular PHSP epitope, represents a true measure of affinity. High-affinity antibody preparations with  $K_a$  ranging from about  $10^9$  to  $10^{12}$  L/mole are preferred for use in immunoassays in which the PHSP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar 30 procedures which ultimately require dissociation of PHSP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J. E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to 35 determine the quality and suitability of such preparations for certain downstream applications. For

example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of PHSP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra,  
5 and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding PHSP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding PHSP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary  
10 to polynucleotides encoding PHSP. Thus, complementary molecules or fragments may be used to modulate PHSP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding PHSP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or  
15 from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding PHSP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding PHSP can be turned off by transforming a cell or tissue with expression  
20 vectors which express high levels of a polynucleotide, or fragment thereof, encoding PHSP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the  
25 vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding PHSP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly,  
30 inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A  
35 complementary sequence or antisense molecule may also be designed to block translation of mRNA

by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding PHSP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding PHSP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nature Biotechnology 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such

therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the  
5 therapeutic effects discussed above. Such pharmaceutical compositions may consist of PHSP, antibodies to PHSP, and mimetics, agonists, antagonists, or inhibitors of PHSP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered  
10 to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

15 In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

20 Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active  
25 compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and  
30 tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene  
35 glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures.



Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol.

5 Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in  
10 aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil,  
15 or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be  
20 permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many  
25 acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

30 After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of PHSP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The  
35 determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example PHSP or fragments thereof, antibodies of PHSP, and agonists, antagonists or inhibitors of PHSP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the  $ED_{50}$  (the dose therapeutically effective in 50% of the population) or  $LD_{50}$  (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the  $LD_{50}/ED_{50}$  ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the  $ED_{50}$  with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1  $\mu\text{g}$  to 100,000  $\mu\text{g}$ , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

#### DIAGNOSTICS

In another embodiment, antibodies which specifically bind PHSP may be used for the diagnosis of disorders characterized by expression of PHSP, or in assays to monitor patients being treated with PHSP or agonists, antagonists, or inhibitors of PHSP. Antibodies useful for diagnostic

purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for PHSP include methods which utilize the antibody and a label to detect PHSP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter  
5 molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring PHSP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of PHSP expression. Normal or standard values for PHSP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to PHSP under conditions suitable  
10 for complex formation. The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of PHSP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding PHSP may be used for  
15 diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of PHSP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of PHSP, and to monitor regulation of PHSP levels during therapeutic intervention.

20 In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding PHSP or closely related molecules may be used to identify nucleic acid sequences which encode PHSP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high,  
25 intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding PHSP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the PHSP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:32-  
30 62 or from genomic sequences including promoters, enhancers, and introns of the PHSP gene.

Means for producing specific hybridization probes for DNAs encoding PHSP include the cloning of polynucleotide sequences encoding PHSP or PHSP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA  
35 polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a

variety of reporter groups, for example, by radionuclides such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding PHSP may be used for the diagnosis of disorders associated with expression of PHSP. Examples of such disorders include, but are not limited to, a cell  
5 proliferative disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall  
10 bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-  
15 candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation,  
20 osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a neuronal disorder, such as akathisia,  
25 Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, bipolar disorder, catatonia, dementia, depression, diabetic neuropathy, Down's syndrome, tardive dyskinesia, dystonias, epilepsy, Huntington's disease, peripheral neuropathy, multiple sclerosis, neurofibromatosis, Parkinson's disease, paranoid psychoses, postherpetic neuralgia, schizophrenia, and Tourette's disorder. The polynucleotide sequences encoding PHSP may be used in Southern or northern analysis, dot blot, or  
30 other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered PHSP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding PHSP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide  
35 sequences encoding PHSP may be labeled by standard methods and added to a fluid or tissue sample

from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding PHSP in the sample  
5 indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of PHSP, a normal or standard profile for expression is established. This may be accomplished by combining  
10 body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding PHSP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples  
15 from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from  
20 successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance  
25 of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding PHSP may involve the use of PCR. These oligomers may be chemically synthesized, generated  
30 enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding PHSP, or a fragment of a polynucleotide complementary to the polynucleotide encoding PHSP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

35 Methods which may also be used to quantify the expression of PHSP include radiolabeling

or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in  
5 various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene  
10 function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci.* 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al.  
15 (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci.* 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding PHSP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to  
20 artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) *Nat Genet.* 15:345-355; Price, C.M. (1993) *Blood Rev.* 7:127-134; and Trask, B.J. (1991) *Trends Genet.* 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome  
25 mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding PHSP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide  
30 sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may  
35 reveal associated markers even if the number or arm of a particular human chromosome is not known.

New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that  
5 area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, PHSP, its catalytic or immunogenic fragments, or  
10 oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between PHSP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds  
15 having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with PHSP, or fragments thereof, and washed. Bound PHSP is then detected by methods well known in the art. Purified PHSP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively,  
20 non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding PHSP specifically compete with a test compound for binding PHSP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PHSP.

25 In additional embodiments, the nucleotide sequences which encode PHSP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding  
30 description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 09/173,482, 09/123,494, 09/152,814, 09/229,005, 60/106,889, 60/109,093,  
35 and 60/113,796, are hereby expressly incorporated by reference.

## EXAMPLES

### I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some  
5 tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized  
and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a  
monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged  
over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either  
isopropanol or sodium acetate and ethanol, or by other routine methods.

10 Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA  
purity. In some cases, RNA was treated with DNase. For most libraries, poly(A<sup>+</sup>) RNA was isolated  
using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN,  
Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was  
isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA  
15 purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA  
libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP  
vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the  
recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units  
20 5.1-6.6). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic  
oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the  
appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000  
bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column  
chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs  
25 were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g.,  
PBLUESCRIPT plasmid (Stratagene), pSPORT1 plasmid (Life Technologies), or pINCY (Incyte  
Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent *E. coli* cells  
including XL1-BLUE, XL1-BLUERF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or  
ELECTROMAX DH10B from Life Technologies.

### 30 II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision, using the UNIZAP vector  
system (Stratagene) or cell lysis. Plasmids were purified using at least one of the following: a Magic  
or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit  
(Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid,  
35 QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid kit from QIAGEN.



Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal  
5 cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a Fluoroskan II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

### III. Sequencing and Analysis

10 cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cyclor or the PTC-200 thermal cyclor (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing  
15 kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing systems (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading  
20 frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the  
25 art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other  
30 parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR).

The polynucleotide sequences were validated by removing vector, linker, and polyA  
35 sequences and by masking ambiguous bases, using algorithms and programs based on BLAST,

dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases, such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Str. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:32-62. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

#### IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel, 1995, *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ database (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

100

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding PHSP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic,

developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation/trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories.

- 5 Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

#### V. Extension of PHSP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:32-62 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this  
10 fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in  
15 hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction  
20 mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and  $\beta$ -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the  
25 alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100  $\mu$ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE  
30 and 0.5  $\mu$ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending  
35 the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulphoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:32-62 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

#### **VI. Labeling and Use of Individual Hybridization Probes**

Hybridization probes derived from SEQ ID NO:32-62 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 µCi of [ $\gamma$ -<sup>32</sup>P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10<sup>7</sup> counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon

membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are compared.

## 5 VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, *supra*.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand  
10 or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

15 Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an  
20 appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

## 25 VIII. Complementary Polynucleotides

Sequences complementary to the PHSP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring PHSP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are  
30 designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of PHSP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the PHSP-encoding transcript.

## IX. Expression of PHSP

35 Expression and purification of PHSP is achieved using bacterial or virus-based expression

systems. For expression of PHSP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory  
5 element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express PHSP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of PHSP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is  
10 replaced with cDNA encoding PHSP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E. K.  
15 et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, PHSP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-  
20 kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from PHSP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman  
25 Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch 10 and 16). Purified PHSP obtained by these methods can be used directly in the following activity assay.

#### X. Demonstration of PHSP Activity

30 PHSP protein kinase is measured by the phosphorylation of a substrate in the presence of gamma-labeled  $^{32}\text{P}$ -ATP. PHSP is incubated with an appropriate substrate and  $^{32}\text{P}$ -ATP in a buffered solution.  $^{32}\text{P}$ -labeled product is separated from free  $^{32}\text{P}$ -ATP by gel electrophoresis or chromatographic procedures, and the incorporated  $^{32}\text{P}$  is quantified by phosphorimage analysis or using a scintillation counter. The amount of  $^{32}\text{P}$  detected is proportional to the activity of PHSP in  
35 this assay. The specific amino acid residue phosphorylated by PHSP may be determined by

phosphoamino acid analysis of the labeled, hydrolyzed protein.

PHSP phosphatase activity is measured by the removal of phosphate from a [ $^{32}\text{P}$ ]-labelled substrate. PHSP is incubated with an appropriate [ $^{32}\text{P}$ ]-labelled substrate in a buffered solution. Reaction products are separated by gel electrophoresis or chromatographic procedures, and the level of  $^{32}\text{P}$  associated with the substrate molecule is quantified by phospho-image analysis or scintillation counting. The difference in  $^{32}\text{P}$  associated with untreated substrate versus PHSP-treated substrate is a measure of phosphatase activity and is proportional to PHSP activity.

#### **XI. Functional Assays**

PHSP function is assessed by expressing the sequences encoding PHSP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10  $\mu\text{g}$  of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2  $\mu\text{g}$  of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate properties, for example, their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M. G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of PHSP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding PHSP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art.

Expression of mRNA encoding PHSP and other genes of interest can be analyzed by northern analysis or microarray techniques.

## **XII. Production of PHSP Specific Antibodies**

PHSP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g.,  
5 Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the PHSP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for  
10 selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase  
15 immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for anti-peptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

## **XIII. Purification of Naturally Occurring PHSP Using Specific Antibodies**

20 Naturally occurring or recombinant PHSP is substantially purified by immunoaffinity chromatography using antibodies specific for PHSP. An immunoaffinity column is constructed by covalently coupling anti-PHSP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

25 Media containing PHSP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PHSP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/PHSP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and PHSP is collected.

## **30 XIV. Identification of Molecules Which Interact with PHSP**

PHSP, or biologically active fragments thereof, are labeled with <sup>125</sup>I Bolton-Hunter reagent. (See, e.g., Bolton et al. (1973) *Biochem. J.* 133:529.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled PHSP, washed, and any wells with labeled PHSP complex are assayed. Data obtained using different concentrations of PHSP are used to  
35 calculate values for the number, affinity, and association of PHSP with the candidate molecules.



Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific  
5   embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

TABLE I

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	32	132240	BMARNOT02	132240H1 and 132240R1 (BMARNOT02), 3254142H1 (OVRTUT01), 1453821X14F1 and 1453821F6 (PENITUT01)
2	33	2180116	SININOT01	2180116H1 and 2180116T6 (SININOT01), 3046645H1 (HEAANOT01), 1918183H1 (PROSNOT06), and 1482405F1 (CORENOT02)
3	34	2197671	SPLNFET02	2197671H1 (SPLNFET02), 666366X22R1 (SCORNOT01), 693783X14 (SYNORAT03), 824265X33F1 (PROSNOT06), 039482R1 and 039482F1 (HUVENOB01), 1453984T6 (PENITUT01), 1663987H1 (BRSTNOT09), and 125901R1 (LUNGNOT01)
4	35	2594943	OVRTUT02	2594943H1 (OVRTUT02), 3617557H1 (EPIPNOT01), 2269005R6 (UTRSNOT02), 1307764F6 (COLNFET02), 1377794F6 (LUNGNOT10), and 1286608H1 (BRAINOT11)
5	36	1513871	PANCTUT01	754239R6 (BRAITUT02), 1513871H1 (PANCTUT01), 2414420F6 (HNT3AZT01), 3291775F6 (BONRFET01), 3821451F6 (BONSTUT01)
6	37	156108	THP1PLB02	156108F1 and 156108H1 (THP1PLB02), 336346R6 (EOSIHET02), 1319528F1 (BLADNOT04), 2375549F6 (ISLTNOT01), SBFA04563F1, SBFA04977F1
7	38	2883243	UTRSTUT05	1342082F6 (COLNTUT03), 1933387T6 (COLNNOT16), 2766460F6 (BRSTNOT12), 2883243H1 (UTRSTUT05), 3524262H1 (ESOGTUN01), 3766487F6 (BRSTNOT24)

TABLE 1 cont.

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
8	39	3173355	UTRSTUT04	1300803F6 and 1300803T6 (BRSTNOT07), 2477542F6 (SMCANOT01), 2477542T6 (SMCANOT01), 2875968H1 (THYRNOT10), 3173355F6 and 3173355H1 (UTRSTUT04), 3290825H1 (BONRFET01), 5192561H1 (OVARBIT06)
9	40	5116906	SMCBUNT01	267517F1 (HNT2NOT01), 263823R1 (HNT2AGT01), 5116906H1 (SMCBUNT01)
10	41	940589	ADRENOT03	029801R6 (SPLNFET01), 940589H1 (ADRENOT03), 1737403T6 (COLNNOT22), 1805477F6 and 1805477T6 (SINTNOT13), 2447613H1 (THP1NOT03), 3408563H1 (PROSTUS08), 3519506H1 (LUNGNOT03), 3637343T6 (LUNGNOT30)
11	42	304421	TESTNOT04	304421H1, 304421X318B2, and 304421X323B2 (TESTNOT04), 2639579F6 (BONTNOT01), 2951859H1 (KIDNFET01)
12	43	1213802	BRSTTUT01	894574R1 (BRSTNOT05), 1213802H1 (BRSTTUT01), 1233414F1 and 1234238H1 (LUNGFET03), 1255782F2 and 1255782T1 (MENITUT03), 1455429F1 (COLNFET02), 1576102T1 (LNODNOT03), 2189267F6 (PROSNOT26), 2748179F6 (LUNGUTUT11), 2831667H1 (TLYMNOT03), 3031229H1 (TLYMNOT05), 3054893H1 (LNODNOT08), 3797030F6 (SPLNNOT12), 3880154H1 (SPLNNOT11), 4852525H1 (TESTNOT10), 5514137H1 (BRADDIR01), 5518378H1 (LIVRDIR01)
13	44	1378134	LUNGNOT10	1378134H1 and 1378134X11 (LUNGNOT10), 2205185F6 (SPLNFET02), 4959694H1 (TLYMNOT05), SAMA00107F1, SAMA00160F1, SAMA0020F1

TABLE 1 cont.

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
14	45	1490070	UCMCL5T01	432218H1 (BRAVUNT02), 1490070H1 (UCMCL5T01), 1535394F1 (SPLNNOT04), 1616509F6 and 1616509T6 (BRAITUT12), 2490845H1 (EOSITXT01), 2723789F6 (LUNGUT01), SAOA00263F1
15	46	1997814	BRSTTUT03	855350R1 (NGANNOT01), 875417R1 (LUNGAST01), 895096R1 (BRSTNOT05), 1271348F1 (TESTTUT02), 1331289F6 (PANCNOT07), 1359243F1 (LUNGNOT12), 1540824T1 (SINTTUT01), 1839828H1 (EOSITXT01), 1997814H1 (BRSTTUT03), 2170638F6 (ENDCNOT03), 3751363F6 (UTRSNOT18)
16	47	2299715	BRSTNOT05	637354R6 and 637354T6 (NEUTGMT01), 1852144F6 (LUNGFET03), 2172576F6 (ENDCNOT03), 2232449F6 (PROSNOT16), 2299715H1 (BRSTNOT05), 2509737X325D2 (CONUTUT01), 2606210F6 (LUNGUT07), 2692024F6 (LUNGNOT23), 2805893F6 (BLADTUT08), 2986160H1 (CARGDIT01), 3085382H1 (HEAONOT03), 3136101F6 and 3136587H1 (SMCCNOT01), 4249977H1 (BRADDIR01)
17	48	209854	SPLNNOT02	209854H1 and 209854T6 (SPLNNOT02), 3152165R6 and 3152165T6 (ADRENON04)
18	49	1384286	BRAITUT08	676123R6 and 676123T6 (CRBLNOT01), 989218X11 and 989218X12 (LVENNOT03), 1384286H1 (BRAITUT08), 3099868H1 (PROSBPT03), 4693167H1 (BRAENOT02)
19	50	1512656	PANCTUT01	322847X5 (EOSIHET02), 1253795T6 (LUNGFET03), 1512656H1 (PANCTUT01), 1561686X303D1 (SPLNNOT04), 2212305H1 (SINTFET03), 2697679H1 (UTRSNOT12), 3205172H1 (PENCNOT03), 5313318H1 (KIDETXS02)

TABLE 1 cont.

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
20	51	2098635	BRAITUT02	1268848T1, 1268848X301F1, and 2157157H1 (BRAINOT09), 2098635H1 and 2098635R6 (BRAITUT02), 2198819F6, 2198819X301D4, 2198819X303D1, 2198819X309B2, and 2198819X309D4 (SPLNFET02), 2784975H2 (BRSTNOT13), 3320340H1 (PROSBPT03)
21	52	2446646	THP1NOT03	000297R6 and 000297X61 (U937NOT01), 2446646H1 (THP1NOT03), 2557274F6 (THYMNOT03)
22	53	2764911	BRSTNOT12	678618T6 and 678618X14 (UTRSNOT02), 2304126R6 (BRSTNOT05), 2764911H1 (BRSTNOT12), 2834475F6 (TLYMNOT03), 2915803F6 (THYMFET03), 3035012F6 (TLYMNOT05), SAFC00027F1, SAFC00254F1, SAFC02376F1, SAFC01609F1
23	54	3013946	MUSCNOT07	673753H1 (CRBLNOT01), 989218X11 and 989218X14 (LVENNOT03), 2821720F6 (ADRETUT06), 3013946F6, 3013946H1, and 3013946T6 (MUSCNOT07), 4693167H1 (BRAENOT02)
24	55	067967	HUVESTB01	067967X92, 067966R1, and 067967H1 (HUVESTB01), SAIA02074F1, SAIA03254F1, SAIA03603F1, and SAIA02259F1
25	56	346275	THYMNOT02	346275H1 (THYMNOT02), 609792X12 (COLNNOT01), SAGA03543F1, SAGA02528F1, and SAGA00285F1
26	57	283746	CARDNOT01	283746H1 and 283746X10 (CARDNOT01), 4903108H1 (TLYMNOT08), 557918X15 (MPHGLPT02), and 2379045F6 (ISLTNOT01)
27	58	2696537	UTRSNOT12	2696537H1 (UTRSNOT12), 3173337F6 (UTRSTUT04), 082658X100 (HUVESTB01), and 603219T6 (BRSTTUT01)

TABLE 1 cont.

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
28	59	551178	BEPINOT01	551178H1 (BEPINOT01), 861522R1 (BRAITUT03), 965838R1 (BRSTNOT05), 1574007F1 and 1574007T1 (LNODNOT03), 1830083T6 and 1831194T6 (THP1AZT01), 3098496H1 (CERVNOT03), 3293481H1 (TLYJINT01)
29	60	619292	PGANNOT01	613165F1 (COLNUT02), 619292H1 and 619292X13 (PGANNOT01)
30	61	2054049	BEPINOT01	1736355F6 (COLNNOT22), 2054049H1 (BEPINOT01), 2379092T6 (ISLTNOT01), 3127284T3 (LUNGTUT12), 3136377F6 (SMCCNOT01), SBMA00545F1, SBMA00827F1, SBMA02930F1, SBMA02853F1
31	62	2843910	DRGLNOT01	036294X71 (HUVENOB01), 066017X102, 068399R1, and 068399X3 (HUVESTB01), 1527276H1 (UCMCL5T01), 1846570T6 (COLNNOT09), 2843910H1 (DRGLNOT01)

TABLE 2

P lypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Homologous sequences	Analytical Methods
1	300	S3 S15 S19 S20 S24 T98 S125 S231 T238 S257 S282 S12 S41 S70 T120 T143 S146 T242	N85 N88 N96	Protein kinase motifs: G161-P256 catalytic tk domain IX: V180-E202	Protein kinase	BLAST PFAM PRINTS
2	147	S85 T38 S90		Calcium-binding repeat motifs: G28-L115	PKC- potentiated inhibitory protein of PP1 (CPI17)	BLAST PRINTS BLOCKS
3	431	T178 S282 T25 S34 S75 S106 S194 S198 T208 T264 S299 S303 S304 S308 T328 S345 S388 T46 S137 S260	N44 N242	PTK signatures: A18-Y283 ATP-binding site: I30-K53, E127-G164 Y196-H219 PK catalytic subdomains: M99-E112, Y134-L152 G181-I191, Y243- A265	Ste20-like protein kinase	BLOCKS PRINTS PROFILESCAN BLAST
4	218	S108 S68 S90 T133 T170 S172 T34 T123 T207		Phosphofructokinase domains: I47, V177-Q195 L148-Y164		PRINTS

TABLE 2 cont.

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Homologous sequences	Analytical Methods
5	474	S14 S89 S98 S132 S472 T22 S26 S62 S66 T204 T320 T345 T359 S427 S443 S94 S128 T211 T336 S443 Y155		Protein kinase family signature: Y144-F425	serine /threonine protein kinase	MOTIFS PFAM BLOCKS PRINTS ProfilesScan BLAST
6	540	S102 S183 S267 T296 T301 S442 S34 S58 S180 S207 S224 T360 S374 S401 S428 S478 T484 Y23	N100 N391 N457 N537	Protein kinase family signature: L18-L287	serine /threonine protein kinase	MOTIFS PFAM BLOCKS PRINTS PROFILESSCAN BLAST
7	454	S57 S69 S130 T203 T212 S338 S420 S91 T101 T220 S271 S295 T315 S359 S381 Y197	N55 N140 N218 N403 N437 N441	SH2 domain: W63-Y138, W354-Y428 PI 3 kinase P85 regulator: K153-G176, A216- N257, R287-N332	phosphatidyl- inositol 3- kinase	PFAM BLOCKS PRINTS BLAST
8	502	S246 T498 T21 S65 S76 T193 T203 S275 S312 S355 T484 S106 T222 S323 T498 Y347	N302 N414	Signal peptide: M1-T21 SH2 domain: V70-E80 ER targeting signal: K499-L502	tyrosine kinase	SigPept BLOCKS MOTIFS BLAST



TABLE 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Homologous sequences	Analytical Methods
9	281	T66 T140 T141 T182 S210	N117 N139	Signal peptide: M1-I76	calcium /calmodulin- dependent protein kinase	PFAM BLAST
10	510	T297 S323 S358 S51 T312 S323 T325 S329 T377 T390 T483 S24 S152 T201 S210 S247 T292 T406 T407	N185 N349 N381 N405	Protein kinase family signature: R52-V261	Serine /threonine protein kinase	PFAM BLOCKS PRINTS MOTIFS BLAST
11	248	S5 S20 S36 T210 T245	N208	Tyrosine specific phosphatase active site: P166-A220 Dual specificity phosphatase: H95-R240	Tyrosine phosphatase or Dual specificity phosphatase	BLAST, MOTIFS BLOCKS, PRINTS PROFILES PFAM

TABLE 2 cont.

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Homologous sequences	Analytical Methods
12	810	S62 S290 T429 S758 T17 T104 S108 T216 S279 T316 S330 T360 S386 T405 S425 S465 T473 S497 T547 T561 T715 S733 S738 S768 S196 S222 S229 S267 T281 T321 T347 S370 T400 T512 S534 T609 S617 S663 S751 T754 T762 Y67	N33		Protein kinase	BLAST, MOTIFS
13	549	S6 T502 T21 T116 S125 S320 T417 S46 S87 T240 S390 S397 S405 S430 S497	N238	ATP/GTP-binding site (p-loop): G58-T65 Protein kinase signature: I176-K199 I292-L304 Y347-L370 F456-L483	Dual specificity tyrosine /serine protein kinase	BLAST, MOTIFS BLOCKS, PRINTS PFAM
14	416	S312 T20 T97 S104 S183 T185 T211 T274 S381 S411 S72 S79 S140 S318 Y53		SH3 domain: A366-D384 N402-E414	PEST phosphatase interacting protein	BLAST, MOTIFS BLOCKS, PRINTS PFAM

TABLE 2 cont.

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Homologous sequences	Analytical Methods
15	425	T34 S233 S234 S25 S107 T144 T198 T250 S251 S258 S282 S300 S324 S345 T390 T51 T133 S365 S383 Y71	N23 N176 N362		SH3 binding protein	BLAST, MOTIFS
16	1135	S77 T187 S259 S554 S815 S9 S17 T59 S112 T124 T222 S264 T319 S324 S326 S550 T572 S625 S681 S682 T688 T689 S706 S720 T931 S958 S978 S999 S255 T309 T351 T543 S550 S624 S632 S726 T811 S898 S1012 S1113 Y321 Y323 Y467	N33 N570 N718 N1067	Protein kinase signature: V31-K54 V149-L161 W129-V182 Tyrosine kinase catalytic site: G190-I200 S214-M236 NIK1-like kinase domain: Y836-R1115	NIK kinase	BLAST, MOTIFS PROFILES BLOCKS, PRINTS PFAM
17	228	T163 S60 T78 T68 S88 S147	N19 N100 N114		Interferon- induced PK regulator (P52rIPK)	BLAST

TABLE 2 cont.

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Homologous sequences	Analytical Methods
18	503	S51 T262 T36 S79 T94 S109 T361 T362 T403 S472 T47 S334 S343 Y17	N313 N333 N360	Protein kinase signature: I20-K43 V132-L144 V195-E217 Protein kinase domain: Y14-V272	calcium /calmodulin- dependent protein kinase II, beta 3 isoform	BLAST, BLOCKS, PRINTS, MOTIFS, PFAM, PROFILES SCAN
19	433	S12 S77 S124 S131 S255 S290 T327 S365 S402 T70 Y88			Choline kinase isolog 384D8_3	BLAST, MOTIFS
20	527	S417 S154 S199 T367 S453 T120 S178 S413 T447 S473	N470	Protein kinase signature: I144-K167 I260-V172 ATP-binding site: Q247-G284 Y318-F341 Protein kinase domain: I138-L427	MAP-related protein kinase	BLAST, BLOCKS MOTIFS, PFAM, PROFILES SCAN

TABLE 2 cont.

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Homologous sequences	Analytical Methods
21	322	S19 S122 T198 T200 T236 S251 T260 S264 T301 S14 S52 T181 T225	N196 N249	Protein kinase signature: L163-I175 ATP-binding site: M150-V187 I224-H247 Protein kinase domain: S32-E316	Protein tyrosine kinase	BLAST, BLOCKS, PRINTS, MOTIFS, PFAM, PROFILES CAN
22	802	S70 T87 S750 T14 T98 S144 T150 S230 S263 T353 T465 T470 S517 S633 T751 S758 T27 T74 T100 T207 S268 S368 S458	N36 N655	Protein kinase signature: L55-K81, L432-K455 ATP-binding site: E160-G197, H232-F255 PTK catalytic domain: H534-F552, C603-H625 Protein kinase domains: F49-F318, L427-L687 Protein kinase C domain: Q319-I382	Ribosomal S6 protein kinase	BLAST, BLOCKS, PRINTS, MOTIFS, PFAM, PROFILES CAN
23	641	S51 T262 S398 S436 S479 T36 S79 T94 S109 T375 T376 T541 S610 T47 S315 S333 S342 S393 S422 S431 S465 S474 S508 Y17	N313 N332 N374	Protein kinase signature: I20-K43 V132-L144 ATP-binding site: Q119-A156 Y191-F214 Protein kinase domain: Y14-V272	Ca2+ /calmodulin dependent protein kinase	BLAST, BLOCKS, PRINTS, MOTIFS, PFAM, PROFILES CAN

TABLE 2 cont.

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Homologous sequences	Analytical Methods
24	588	S106 T155 S359 T388 T456 T531 T4, S58 S108 T126 S132 T279 S350 S436 S469 S508 S537 Y32	N63 N130 N574	Protein kinase catalytic domain: Y209-S445, F495-I522 ATP-binding site: I215-K238 STK core catalytic motif: I331-L343	Protein kinase Dyrk2	MOTIFS PFAM BLOCKS PRINTS BLAST
25	389	S31 T301 S56 S96 S134 T149 S186 S201 S283 S358 S375 Y148 Y165	N257 N343 N364	Protein kinase catalytic domain: E73-I311 STK core catalytic motif: I172-Y184 PTK core domain: D152-D208	CaM-like protein kinase	BLAST PFAM MOTIFS BLOCKS PRINTS PROFILESCAN
26	343	S68 S81 S137 S184 T219 S276 S297 T29 T125 Y86 Y211	N332	EF hand calcium-binding signature: D176-L188	protein phosphatase 2A (PR72)	BLAST MOTIFS BLOCKS
27	184	S36 T105 S40 S70 T117 Y50	N62	Tyrosine phosphatase active site domain: L63-V118	MAP kinase phosphatase (X17C)	BLAST PROFILESCAN BLOCKS PRINTS MOTIFS

TABLE 2 cont.

P lypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Homologous sequences	Analytical Methods
28	367	S10 S21 S44 S103 T116 T267 T309 S191 S213 S218 S256 T305 S352 Y159 Y344	N16 N17		protein phosphatase 2A, A-subunit	BLAST
29	118	S34 S84	N43	Signal peptide: M1-A27 PDZ domain: H8-S73	tyrosine phosphatase	SPScan PFAM BLAST
30	356	S9 S94 T209 T220 S259 S337 S5 S26 S75 S121 T154 S282 S332 S339 Y15 Y84	N333	tyrosine-specific protein phosphatase active site: I108-K164	tyrosine phosphatase (myotubularin)	PROFILES MOTIFS BLOCKS PRINTS BLAST
31	453	S38 S73 S119 S131 S193 S200 T236 S293 S341 T379 T124 S173 T214 S252 T256 S282 S302 S313 S391 S397	N43 N67 N357	protein phosphatase 2A p55 subunit: P10-K451	protein phosphatase 2A p55 regulatory subunit, alpha isoform	PFAM MOTIFS BLOCKS PRINTS BLAST

TABLE 3

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
32	Hematopoietic/Immune (0.333) Reproductive (0.333)	Cell proliferation (0.500) Inflammation (0.333)	PBLUESCRIPT
33	Nervous (0.216) Reproductive (0.235) Cardiovascular (0.118)	Cell proliferation (0.530) Inflammation (0.352)	pINCY
34	Reproductive (0.293) Gastrointestinal (0.192)	Cell proliferation (0.641) Inflammation (0.335)	pINCY
35	Reproductive (0.284) Nervous (0.210) Cardiovascular (0.1213)	Cell proliferation (0.729) Inflammation (0.272)	pINCY
36	Nervous (0.529) Developmental (0.118) Gastrointestinal (0.118)	Cell proliferation (0.588) Neurological (0.118) Inflammation (0.118)	pINCY
37	Hematopoietic/Immune (0.268) Reproductive (0.244) Nervous (0.122)	Inflammation (0.488) Cell Proliferative (0.415)	PBLUESCRIPT
38	Reproductive (0.400) Hematopoietic/Immune (0.160) Nervous (0.160)	Cell proliferation (0.600) Inflammation (0.320)	pINCY
39	Cardiovascular (0.312) Reproductive (0.312) Developmental (0.188)	Cell proliferation (0.938) Inflammation (0.125)	pINCY
40	Nervous (0.400) Gastrointestinal (0.267) Developmental (0.133)	Cell proliferation (0.733) Neurological (0.133) Inflammation (0.133)	pINCY
41	Gastrointestinal (0.267) Nervous (0.233) Reproductive (0.167)	Inflammation (0.533) Cell proliferation (0.534)	pSPORT1



Table 3 cont.

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
42	Musculoskeletal (0.500) Developmental (0.167) Gastrointestinal (0.167)	Cancer (0.834) Inflammation (0.167)	PBLUESCRIPT
43	Reproductive (0.240) Nervous (0.151) Gastrointestinal (0.135)	Cell proliferation (0.536) Inflammation (0.417)	pSPORT1
44	Hematopoietic/Immune (0.278) Nervous (0.222) Dermatologic (0.111)	Cell proliferation (0.444) Inflammation (0.389)	pINCY
45	Hematopoietic/Immune (0.500) Gastrointestinal (0.125) Nervous (0.125)	Inflammation (0.500) Cell proliferative (0.500)	PBLUESCRIPT
46	Nervous (0.220) Reproductive (0.213) Hematopoietic/Immune (0.140)	Cell proliferation (0.573) Inflammation (0.380)	pSPORT1
47	Hematopoietic/Immune (0.190) Gastrointestinal (0.165) Nervous (0.139)	Cell proliferation (0.582) Inflammation (0.354)	pSPORT1

Table 3 cont.

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
48	Nervous (0.333) Reproductive (0.333) Hematopoietic/Immune (0.111)	Cancer (0.444) Inflammation (0.222) Neurological (0.111)	PBLUESCRIPT
49	Nervous (0.724) Cardiovascular (0.103)	Inflammation (0.276) Cancer (0.241) Neurological (0.172)	pINCY
50	Reproductive (0.235) Hematopoietic/Immune (0.188) Gastrointestinal (0.129)	Cancer (0.447) Inflammation (0.282) Fetal (0.153)	pINCY
51	Nervous (0.368) Developmental (0.158) Gastrointestinal (0.105)	Cancer (0.368) Fetal (0.211) Inflammation (0.105)	pSPORT1
52	Cardiovascular (0.312) Hematopoietic/Immune (0.312) Reproductive (0.158)	Fetal (0.688) Cancer (0.421) Inflammation (0.125)	pINCY
53	Reproductive (0.412) Nervous (0.235) Developmental (0.118)	Cancer (0.471) Fetal (0.235) Inflammation (0.235)	pINCY
54	Nervous (0.714) Cardiovascular (0.107)	Cancer (0.250) Inflammation (0.250) Neurological (0.179)	pINCY

Table 3 cont.

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	PBLUESCRIPT
55	Reproductive (0.533) Nervous (0.133)	Cell proliferation (0.601) Inflammation (0.270)	PBLUESCRIPT
56	Hematopoietic/Immune (0.278) Nervous (0.222) Reproductive (0.154)	Cell proliferation (0.388) Inflammation (0.333) Neurological (0.111)	PBLUESCRIPT
57	Hematopoietic/Immune (0.211) Cardiovascular (0.193) Nervous (0.175)	Cell proliferation (0.474) Inflammation (0.491)	PBLUESCRIPT
58	Reproductive (0.286) Cardiovascular (0.229) Musculoskeletal (0.143)	Cell proliferation (0.715) Inflammation (0.200)	PINCY
59	Reproductive (0.253) Gastrointestinal (0.211) Nervous (0.147)	Cancer and Cell proliferation (0.684) Inflammation and Immune Response (0.242)	PSPORT1
60	Nervous (0.667) Reproductive (0.333)	Cancer (1.000)	PSPORT1
61	Reproductive (0.357) Cardiovascular (0.179) Nervous (0.125)	Cancer and Cell proliferation (0.642) Inflammation and Immune Response (0.232)	PSPORT1
62	Nervous (0.228) Reproductive (0.175) Cardiovascular (0.158) Hematopoietic/Immune (0.158)	Cancer (0.368) Inflammation and Immune Response (0.263) Fetal (0.211)	PINCY

TABLE 4

Polynucleotide SEQ ID NO:	Library	Library Comment
32	BMARNOT02	Library was constructed using RNA isolated from the bone marrow of 24 male and female Caucasian donors, 16 to 70 years old.
33	SININOT01	Library was constructed using RNA isolated from ileum tissue removed from the small intestine of a 4-year-old Caucasian female, who died from a closed head injury. Patient history included jaundice as a baby. Previous surgeries included a double hernia repair
34	SPLNFET02	Library was constructed using RNA isolated from spleen tissue removed from a Caucasian male fetus, who died at 23 weeks' gestation from premature birth. Family history included diabetes.
35	OVARTUT02	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 51-year-old Caucasian female during an exploratory laparotomy, total abdominal hysterectomy, salpingo-oophorectomy, and an incidental appendectomy. Pathology indicated mucinous cystadenoma presenting as a multiloculated neoplasm involving the entire left ovary. The right ovary contained a follicular cyst and a hemorrhagic corpus luteum. The uterus showed proliferative endometrium and a single intramural leiomyoma. The peritoneal biopsy indicated benign glandular inclusions consistent with endosalpingiosis. The patient presented with abnormal weight gain and ascites. Patient history included depressive disorder, joint pain, allergies, alcohol use, and a normal delivery. Family history included atherosclerotic coronary artery disease, benign hypertension, breast cancer and uterine cancer.

TABLE 4 cont.

Polynucleotide SEQ ID NO:	Library	Library Comment
36	PANCTUT01	Library was constructed using RNA isolated from pancreatic tumor tissue removed from a 65-year-old Caucasian female during radical subtotal pancreatectomy. Pathology indicated an invasive grade 2 adenocarcinoma. Patient history included type II diabetes, osteoarthritis, cardiovascular disease, and benign neoplasm in the large bowel. Previous surgeries included a total splenectomy, cholecystectomy, and abdominal hysterectomy. Family history included cardiovascular disease, type II diabetes, and stomach cancer.
37	SMCBUNT01	Library was constructed using RNA isolated from bronchial smooth muscle c ll tissue removed from a 21-year-old Caucasian male.
38	UTRSTUT05	Library was constructed using RNA isolated from uterine tumor tissue removed from a 41-year-old Caucasian female during a vaginal hysterectomy with dilation and curettage. Pathology indicated uterine leiomyoma. The endometrium was secretory and contained fragments of endometrial polyps. Benign endo- and ectocervical mucosa were identified in the endocervix. Patient history included a ventral hernia and a benign ovarian neoplasm.
39	UTRSTUT04	Library was constructed using RNA isolated from uterine tumor tissue removed from a 34-year-old Caucasian female during a hysteroscopy and an exploratory laparotomy with dilation and curettage. Pathology indicated an endometrial polyp, subserosal leiomyoma, and fragments of leiomyoma. Family history included hyperlipidemia, depressive disorder, benign hypertension, cerebrovascular disease, arteriosclerotic cardiovascular disease, and type II diabetes.

TABLE 4 cont.

Polynucleotide SEQ ID NO:	Library	Library Comment
40	SMCBUNT01	Library was constructed using RNA isolated from bronchial smooth muscle cell tissue removed from a 21-year-old Caucasian male.
41	ADRENOT03	Library was constructed using RNA isolated from the adrenal tissue of a 17-year-old Caucasian male, who died from cerebral anoxia.
42	TESTNOT04	Library was constructed using RNA isolated from testicular tissue removed from a 37-year-old Caucasian male who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.
43	BRSTTUT01	Library was constructed using RNA isolated from breast tumor tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated invasive grade 4 mammary adenocarcinoma of mixed lobular and ductal type, extensively involving the left breast. The tumor was identified in the deep dermis near the lactiferous ducts with extracapsular extension. Seven mid and low and five high axillary lymph nodes were positive for tumor. Proliferative fibrocystic changes were characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Patient history included atrial tachycardia, blood in the stool, and a benign breast neoplasm. Family history included benign hypertension, atherosclerotic coronary artery disease, cerebrovascular disease, and depressive disorder.
44	LUNGNOT10	Library was constructed using RNA isolated from the lung tissue of a Caucasian male fetus who died at 23 weeks' gestation.
45	UCMCL5T01	Library was constructed using RNA isolated from mononuclear cells obtained from the umbilical cord blood of 12 individuals. The cells were cultured for 12 days with IL-5 before RNA was isolated from the pooled lysates.

TABLE 4 cont.

Polynucleotide SEQ ID NO:	Library	Library Comment
46	BRSTTUT03	Library was constructed using RNA isolated from breast tumor tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated multicentric invasive grade 4 lobular carcinoma. The mass was identified in the upper outer quadrant, and three separate nodules were found in the lower outer quadrant of the left breast. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular disease, coronary artery aneurysm, breast cancer, prostate cancer, atherosclerotic coronary artery disease, and type I diabetes.
47	BRSTNOT05	Library was constructed using RNA isolated from breast tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated multicentric invasive grade 4 lobular carcinoma. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular and cardiovascular disease, breast and prostate cancer, and type I diabetes.

TABLE 4 cont.

Polynucleotide SEQ ID NO:	Library	Library Comment
48	SPLNNOT02	The library was constructed using RNA isolated from the spleen tissue of a 29-year-old Caucasian male, who died from head trauma. Serologies were positive for cytomegalovirus (CMV). Patient history included alcohol, marijuana, and tobacco use.
49	BRAITUT08	The library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 47-year-old Caucasian male during excision of cerebral meningeal tissue. Pathology indicated grade 4 fibrillary astrocytoma with focal tumoral radionecrosis. Patient history included cerebrovascular disease, deficiency anemia, hyperlipidemia, epilepsy, and tobacco use. Family history included cerebrovascular disease and a malignant prostate neoplasm.
50	PANCTUT01	The library was constructed using RNA isolated from pancreatic tumor tissue removed from a 65-year-old Caucasian female during radical subtotal pancreatectomy. Pathology indicated an invasive grade 2 adenocarcinoma. Patient history included type II diabetes, osteoarthritis, cardiovascular disease, benign neoplasm in the large bowel, and a cataract. Previous surgeries included a total splenectomy, cholecystectomy, and abdominal hysterectomy. Family history included cardiovascular disease, type II diabetes, and stomach cancer.
51	BRAITUT02	The library was constructed using RNA isolated from brain tumor tissue removed from the frontal lobe of a 58-year-old Caucasian male during excision of a cerebral meningeal lesion. Pathology indicated a grade 2 metastatic hypernephroma. Patient history included a grade 2 renal cell carcinoma, insomnia, and chronic airway obstruction. Family history included a malignant neoplasm of the kidney.



TABLE 4 cont.

Polynucleotide SEQ ID NO:	Library	Library Comment
52	THP1NOT03	The library was constructed using RNA isolated from untreated THP-1 cells. THP-1 (ATCC TIB 202) is a human promonocyte line derived from the peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (ref: Int. J. Cancer (1980) 26:171).
53	BRSTNOT12	The library was constructed using RNA isolated from diseased breast tissue removed from a 32-year-old Caucasian female during a bilateral reduction mammoplasty. Pathology indicated nonproliferative fibrocystic disease. Family history included cardiovascular disease.
54	MUSCNOT07	The library was constructed using RNA isolated from muscle tissue removed from the forearm of a 38-year-old Caucasian female during a soft tissue excision. Pathology for the associated tumor tissue indicated intramuscular hemangioma. Family history included breast cancer, benign hypertension, cerebrovascular disease, colon cancer, and type II diabetes.
55	HUVESTB01	Library was constructed using RNA isolated from shear-stressed HUV-EC-C (ATCC CRL 1730) cells. HUV-EC-C is an endothelial cell line derived from the vein of a normal human umbilical cord (ref:PNAS 81:6413).
56	THYMNOT02	Library was constructed using polyA RNA isolated from thymus tissue removed from a 3-year-old Caucasian male, who died from drowning.
57	CARDNOT01	Library was constructed using RNA isolated from the cardiac muscle of a 65-year-old Caucasian male, who died from a self-inflicted gunshot wound.

TABLE 4 cont.

Polynucleotide SEQ ID NO:	Library	Library Comment
58	UTRSNOT12	Library was constructed using RNA isolated from uterine myometrial tissue removed from a 41-year-old Caucasian female during a vaginal hysterectomy with a dilatation and curettage. The endometrium was secretory and contained fragments of endometrial polyps. Benign endo- and ectocervical mucosa were identified in the endocervix. Pathology for the associated tumor tissue indicated uterine leiomyoma. The patient presented with an unspecified menstrual disorder. Patient history included ventral hernia, normal delivery, a benign ovarian neoplasm, and tobacco abuse. Previous surgeries included a bilateral destruction of fallopian tubes, removal of a solitary ovary, and an exploratory laparotomy.
59	BEPINOT01	Library was constructed using RNA isolated from a bronchial epithelium primary cell line derived from a 54-year-old Caucasian male.
60	PGANNOT01	Library was constructed using RNA isolated from paraganglionic tumor tissue removed from the intra-abdominal region of a 46-year-old Caucasian male during exploratory laparotomy. Pathology indicated a benign paraganglioma and association with a grade 2 renal cell carcinoma, clear cell type.
61	BEPINOT01	Library was constructed using RNA isolated from a bronchial epithelium primary cell line derived from a 54-year-old Caucasian male.
62	DRGLNOT01	Library was constructed using RNA isolated from dorsal root ganglion tissue removed from the low thoracic/high lumbar region of a 32-year-old Caucasian male who died from acute pulmonary edema and bronchopneumonia, bilateral pleural and pericardial effusions, and malignant lymphoma (natural killer cell type). Patient history included probable cytomegalovirus infection, hepatic congestion and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, and Bell's palsy.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value= 1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value= 1.0E-8 or less Full Length sequences: fastx score= 100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Atwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score= 1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score= 10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality scores>GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof.
- 5 2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.
3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.
4. An isolated and purified polynucleotide variant having at least 80% polynucleotide sequence identity to the polynucleotide of claim 3.
- 10 5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.
6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.
7. A method for detecting a polynucleotide, the method comprising the steps of:
  - 15 (a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and
  - (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.
8. The method of claim 7 further comprising amplifying the polynucleotide prior to  
20 hybridization.
9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:32-62 and fragments thereof.
10. An isolated and purified polynucleotide variant having at least 80% polynucleotide sequence identity to the polynucleotide of claim 9.
- 25 11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.
12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.
13. A host cell comprising the expression vector of claim 12.
- 30 14. A method for producing a polypeptide, the method comprising the steps of:
  - a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and
  - b) recovering the polypeptide from the host cell culture.
15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction  
35 with a suitable pharmaceutical carrier.
16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.
18. A purified antagonist of the polypeptide of claim 1.
19. A method for treating or preventing a disorder associated with decreased expression or activity of PHSP, the method comprising administering to a subject in need of such treatment an  
5 effective amount of the pharmaceutical composition of claim 15.
20. A method for treating or preventing a disorder associated with increased expression or activity of PHSP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

## SEQUENCE LISTING

<110> INYCTE PHARMACEUTICALS, INC.  
 HILLMAN, Jennifer L.  
 LAL, Preeti  
 TANG, Y. Tom  
 CORLEY, Neil C.  
 GUEGLER, Karl J.  
 BAUGHN, Mariah R.  
 PATTERSON, Chandra  
 BANDMAN, Olga  
 AU-YOUNG, Janice  
 GORGONE, Gina A.  
 YUE, Henry  
 AZIMZAI, Yalda  
 REDDY, Roopa  
 LU, Dyung Aina M.  
 SHIH, Leo L.

<120> PHOSPHORYLATION EFFECTORS

<130> PF-0565 PCT

<140> To Be Assigned

<141> Herewith

<150> 09/123,494; unassigned; 09/152,814; unassigned; 09/173,482;  
 unassigned;60/106,889; 60/109,093; 60/113,796;

<151> 1998-07-28; 1998-07-28; 1998-09-14; 1998-09-14; 1998-10-14;  
 1998-10-14;1998-11-03; 1998-11-19; 1998-12-22

<160> 61

<170> PERL Program

<210> 1

<211> 300

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 132240

<400> 1

Met	Glu	Ser	Pro	Leu	Glu	Ser	Gln	Pro	Leu	Asp	Ser	Asp	Arg	Ser
1				5					10					15
Ile	Lys	Glu	Ser	Ser	Phe	Glu	Glu	Ser	Asn	Ile	Glu	Asp	Pro	Leu
				20					25					30
Ile	Val	Thr	Pro	Asp	Cys	Gln	Glu	Lys	Thr	Ser	Pro	Lys	Gly	Val
				35					40					45
Glu	Asn	Pro	Ala	Val	Gln	Glu	Ser	Asn	Gln	Lys	Met	Leu	Gly	Pro
				50					55					60
Pro	Leu	Glu	Val	Leu	Lys	Thr	Leu	Ala	Ser	Lys	Arg	Asn	Ala	Val
				65					70					75

```

Ala Phe Arg Ser Phe Asn Ser His Ile Asn Ala Ser Asn Asn Ser
      80      85      90
Glu Pro Ser Arg Met Asn Met Thr Ser Leu Asp Ala Met Asp Ile
      95     100     105
Ser Cys Ala Tyr Ser Gly Ser Tyr Pro Met Ala Ile Thr Pro Thr
     110     115     120
Gln Lys Arg Arg Ser Cys Met Pro His Gln Thr Pro Asn Gln Ile
     125     130     135
Lys Ser Gly Thr Pro Tyr Arg Thr Pro Lys Ser Val Arg Arg Gly
     140     145     150
Val Ala Pro Val Asp Asp Gly Arg Ile Leu Gly Thr Pro Asp Tyr
     155     160     165
Leu Ala Pro Glu Leu Leu Leu Gly Arg Ala His Gly Pro Ala Val
     170     175     180
Asp Trp Trp Ala Leu Gly Val Cys Leu Phe Glu Phe Leu Thr Gly
     185     190     195
Ile Pro Pro Phe Asn Asp Glu Thr Pro Gln Gln Val Phe Gln Asn
     200     205     210
Ile Leu Lys Arg Asp Ile Pro Trp Pro Glu Gly Glu Glu Lys Leu
     215     220     225
Ser Asp Asn Ala Gln Ser Ala Val Glu Ile Leu Leu Thr Ile Asp
     230     235     240
Asp Thr Lys Arg Ala Gly Met Lys Glu Leu Lys Arg His Pro Leu
     245     250     255
Phe Ser Asp Val Asp Trp Glu Asn Leu Gln His Gln Thr Met Pro
     260     265     270
Phe Ile Pro Gln Pro Asp Asp Glu Thr Asp Thr Ser Tyr Phe Glu
     275     280     285
Ala Arg Asn Thr Ala Gln His Leu Thr Val Ser Gly Phe Ser Leu
     290     295     300

```

&lt;210&gt; 2

&lt;211&gt; 147

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2180116

&lt;400&gt; 2

```

Met Ala Ala Gln Arg Leu Gly Lys Arg Val Leu Ser Lys Leu Gln
  1      5      10      15
Ser Pro Ser Arg Ala Arg Gly Pro Gly Gly Ser Pro Gly Gly Met
     20     25     30
Gln Lys Arg His Ala Arg Val Thr Val Lys Tyr Asp Arg Arg Glu
     35     40     45
Leu Gln Arg Arg Leu Asp Val Glu Lys Trp Ile Asp Gly Arg Leu
     50     55     60
Glu Glu Leu Tyr Arg Gly Met Glu Ala Asp Met Pro Asp Glu Ile
     65     70     75
Asn Ile Asp Glu Leu Leu Glu Leu Glu Ser Glu Glu Glu Arg Ser
     80     85     90
Arg Lys Ile Gln Gly Leu Leu Lys Ser Cys Gly Lys Pro Val Glu
     95    100    105

```



```

Asp Phe Ile Gln Glu Leu Leu Ala Lys Leu Gln Gly Leu His Arg
      110                      115                      120
Gln Pro Gly Leu Arg Gln Pro Ser Pro Ser His Asp Gly Ser Leu
      125                      130                      135
Ser Pro Leu Gln Asp Arg Ala Arg Thr Ala His Pro
      140                      145

```

&lt;210&gt; 3

&lt;211&gt; 431

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2197671

&lt;400&gt; 3

```

Met Ala His Ser Pro Val Gln Ser Gly Leu Pro Gly Met Gln Asn
  1                      5                      10                      15
Leu Lys Ala Asp Pro Glu Glu Leu Phe Thr Lys Leu Glu Lys Ile
      20                      25                      30
Gly Lys Gly Ser Phe Gly Glu Val Phe Lys Gly Ile Asp Asn Arg
      35                      40                      45
Thr Gln Lys Val Val Ala Ile Lys Ile Ile Asp Leu Glu Glu Ala
      50                      55                      60
Glu Asp Glu Ile Glu Asp Ile Gln Gln Glu Ile Thr Val Leu Ser
      65                      70                      75
Gln Cys Asp Ser Pro Tyr Val Thr Lys Tyr Tyr Gly Ser Tyr Leu
      80                      85                      90
Lys Asp Thr Lys Leu Trp Ile Ile Met Glu Tyr Leu Gly Gly Gly
      95                      100                     105
Ser Ala Leu Asp Leu Leu Glu Pro Gly Arg Leu Asp Glu Thr Gln
      110                     115                     120
Ile Ala Thr Ile Leu Arg Glu Ile Leu Lys Gly Leu Asp Tyr Leu
      125                     130                     135
His Ser Glu Lys Lys Ile His Arg Asp Ile Lys Ala Ala Asn Val
      140                     145                     150
Leu Leu Ser Glu His Gly Glu Val Lys Leu Ala Asp Phe Gly Val
      155                     160                     165
Ala Gly Gln Leu Thr Asp Thr Gln Ile Lys Arg Asn Thr Phe Val
      170                     175                     180
Gly Thr Pro Phe Trp Met Ala Pro Glu Val Ile Lys Gln Ser Ala
      185                     190                     195
Tyr Asp Ser Lys Ala Asp Ile Trp Ser Leu Gly Ile Thr Ala Ile
      200                     205                     210
Glu Leu Ala Arg Gly Glu Pro Pro His Ser Glu Leu His Pro Met
      215                     220                     225
Lys Val Leu Phe Leu Ile Pro Lys Asn Asn Pro Pro Thr Leu Glu
      230                     235                     240
Gly Asn Tyr Ser Lys Pro Leu Lys Glu Phe Val Glu Ala Cys Leu
      245                     250                     255
Asn Lys Glu Pro Ser Phe Arg Pro Thr Ala Lys Glu Leu Leu Lys
      260                     265                     270
His Lys Phe Ile Leu Arg Asn Ala Lys Lys Thr Ser Tyr Leu Thr
      275                     280                     285

```

Glu	Leu	Ile	Asp	Arg	Tyr	Lys	Arg	Trp	Lys	Ala	Glu	Gln	Ser	His
				290					295					300
Asp	Asp	Ser	Ser	Ser	Glu	Asp	Ser	Asp	Ala	Glu	Thr	Asp	Gly	Gln
				305					310					315
Ala	Ser	Gly	Gly	Ser	Asp	Ser	Gly	Asp	Trp	Ile	Phe	Thr	Ile	Arg
				320					325					330
Glu	Lys	Asp	Pro	Lys	Asn	Leu	Glu	Asn	Gly	Ala	Leu	Gln	Pro	Ser
				335					340					345
Asp	Leu	Asp	Arg	Asn	Lys	Met	Lys	Asp	Ile	Pro	Lys	Arg	Pro	Phe
				350					355					360
Ser	Gln	Cys	Leu	Ser	Thr	Ile	Ile	Ser	Pro	Leu	Phe	Ala	Glu	Leu
				365					370					375
Lys	Glu	Lys	Ser	Gln	Ala	Cys	Gly	Gly	Asn	Leu	Gly	Ser	Ile	Glu
				380					385					390
Glu	Leu	Arg	Gly	Ala	Ile	Tyr	Leu	Ala	Glu	Glu	Ala	Cys	Pro	Gly
				395					400					405
Ile	Ser	Asp	Thr	Met	Val	Ala	Gln	Leu	Val	Gln	Arg	Leu	Gln	Arg
				410					415					420
Tyr	Ser	Leu	Ser	Gly	Gly	Gly	Thr	Ser	Ser	His				
				425					430					

&lt;210&gt; 4

&lt;211&gt; 218

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2594943

&lt;400&gt; 4

Met	Asn	Cys	Arg	Ser	Glu	Val	Leu	Glu	Val	Ser	Val	Glu	Gly	Arg
1				5					10					15
Gln	Val	Glu	Glu	Ala	Met	Leu	Ala	Val	Leu	His	Thr	Val	Leu	Leu
				20					25					30
His	Arg	Ser	Thr	Gly	Lys	Phe	His	Tyr	Lys	Lys	Glu	Gly	Thr	Tyr
				35					40					45
Ser	Ile	Gly	Thr	Val	Gly	Thr	Gln	Asp	Val	Asp	Cys	Asp	Phe	Ile
				50					55					60
Asp	Phe	Thr	Tyr	Val	Arg	Val	Ser	Ser	Glu	Glu	Leu	Asp	Arg	Ala
				65					70					75
Leu	Arg	Lys	Val	Val	Gly	Glu	Phe	Lys	Asp	Ala	Leu	Arg	Asn	Ser
				80					85					90
Gly	Gly	Asp	Gly	Leu	Gly	Gln	Met	Ser	Leu	Glu	Phe	Tyr	Gln	Lys
				95					100					105
Lys	Lys	Ser	Arg	Trp	Pro	Phe	Ser	Asp	Glu	Cys	Ile	Pro	Trp	Glu
				110					115					120
Val	Trp	Thr	Val	Lys	Val	His	Val	Val	Ala	Leu	Ala	Thr	Glu	Gln
				125					130					135
Glu	Arg	Gln	Ile	Cys	Arg	Glu	Lys	Val	Gly	Glu	Lys	Leu	Cys	Glu
				140					145					150
Lys	Ile	Ile	Asn	Ile	Val	Glu	Val	Met	Asn	Arg	His	Glu	Tyr	Leu
				155					160					165
Pro	Lys	Met	Pro	Thr	Gln	Ser	Glu	Val	Asp	Asn	Val	Phe	Asp	Thr
				170					175					180

Gly Leu Arg Asp Val Gln Pro Tyr Leu Tyr Lys Ile Ser Phe Gln  
 185 190 195  
 Ile Thr Asp Ala Leu Gly Thr Ser Val Thr Thr Thr Met Arg Arg  
 200 205 210  
 Leu Ile Lys Asp Thr Leu Ala Leu  
 215

<210> 5  
 <211> 474  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone Number: 1513871

<400> 5  
 Met Ile Met Asn Lys Met Lys Asn Phe Lys Arg Arg Phe Ser Leu  
 1 5 10 15  
 Ser Val Pro Arg Thr Glu Thr Ile Glu Glu Ser Leu Ala Glu Phe  
 20 25 30  
 Thr Glu Gln Phe Asn Gln Leu His Asn Arg Arg Asn Glu Asn Leu  
 35 40 45  
 Gln Leu Gly Pro Leu Gly Arg Asp Pro Pro Gln Glu Cys Ser Thr  
 50 55 60  
 Phe Ser Pro Thr Asp Ser Gly Glu Glu Pro Gly Gln Leu Ser Pro  
 65 70 75  
 Gly Val Gln Phe Gln Arg Arg Gln Asn Gln Arg Arg Phe Ser Met  
 80 85 90  
 Glu Asp Val Ser Lys Arg Leu Ser Leu Pro Met Asp Ile Arg Leu  
 95 100 105  
 Pro Gln Glu Phe Leu Gln Lys Leu Gln Met Glu Ser Pro Asp Leu  
 110 115 120  
 Pro Lys Pro Leu Ser Arg Met Ser Arg Arg Ala Ser Leu Ser Asp  
 125 130 135  
 Ile Gly Phe Gly Lys Leu Glu Thr Tyr Val Lys Leu Asp Lys Leu  
 140 145 150  
 Gly Glu Gly Thr Tyr Ala Thr Val Phe Lys Gly Arg Ser Lys Leu  
 155 160 165  
 Thr Glu Asn Leu Val Ala Leu Lys Glu Ile Arg Leu Glu His Glu  
 170 175 180  
 Glu Gly Ala Pro Cys Thr Ala Ile Arg Glu Val Ser Leu Leu Lys  
 185 190 195  
 Asn Leu Lys His Ala Asn Ile Val Thr Leu His Asp Leu Ile His  
 200 205 210  
 Thr Asp Arg Ser Leu Thr Leu Val Phe Glu Tyr Leu Asp Ser Asp  
 215 220 225  
 Leu Lys Gln Tyr Leu Asp His Cys Gly Asn Leu Met Ser Met His  
 230 235 240  
 Asn Val Lys Ile Phe Met Phe Gln Leu Leu Arg Gly Leu Ala Tyr  
 245 250 255  
 Cys His His Arg Lys Ile Leu His Arg Asp Leu Lys Pro Gln Asn  
 260 265 270  
 Leu Leu Ile Asn Glu Arg Gly Glu Leu Lys Leu Ala Asp Phe Gly  
 275 280 285

```

Leu Ala Arg Ala Lys Ser Val Pro Thr Lys Thr Tyr Ser Asn Glu
      290      295      300
Val Val Thr Leu Trp Tyr Arg Pro Pro Asp Val Leu Leu Gly Ser
      305      310      315
Thr Glu Tyr Ser Thr Pro Ile Asp Met Trp Gly Val Gly Cys Ile
      320      325      330
His Tyr Glu Met Ala Thr Gly Arg Pro Leu Phe Pro Gly Ser Thr
      335      340      345
Val Lys Glu Glu Leu His Leu Ile Phe Arg Leu Leu Gly Thr Pro
      350      355      360
Thr Glu Glu Thr Trp Pro Gly Val Thr Ala Phe Ser Glu Phe Arg
      365      370      375
Thr Tyr Ser Phe Pro Cys Tyr Leu Pro Gln Pro Leu Ile Asn His
      380      385      390
Ala Pro Arg Leu Asp Thr Asp Gly Ile His Leu Leu Ser Ser Leu
      395      400      405
Leu Leu Tyr Glu Ser Lys Ser Arg Met Ser Ala Glu Ala Ala Leu
      410      415      420
Ser His Ser Tyr Phe Arg Ser Leu Gly Glu Arg Val His Gln Leu
      425      430      435
Glu Asp Thr Ala Ser Ile Phe Ser Leu Lys Glu Ile Gln Leu Gln
      440      445      450
Lys Asp Pro Gly Tyr Arg Gly Leu Ala Phe Gln Gln Pro Gly Arg
      455      460      465
Gly Lys Asn Arg Arg Gln Ser Ile Phe
      470

```

&lt;210&gt; 6

&lt;211&gt; 540

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 156108

&lt;400&gt; 6

```

Met Asn Gly Glu Ala Ile Cys Ser Ala Leu Pro Thr Ile Pro Tyr
  1      5      10      15
His Lys Leu Ala Asp Leu Arg Tyr Leu Ser Arg Gly Ala Ser Gly
      20      25      30
Thr Val Ser Ser Ala Arg His Ala Asp Trp Arg Val Gln Val Ala
      35      40      45
Val Lys His Leu His Ile His Thr Pro Leu Leu Asp Ser Glu Arg
      50      55      60
Lys Asp Val Leu Arg Glu Ala Glu Ile Leu His Lys Ala Arg Phe
      65      70      75
Ser Tyr Ile Leu Pro Ile Leu Gly Ile Cys Asn Glu Pro Glu Phe
      80      85      90
Leu Gly Ile Val Thr Glu Tyr Met Pro Asn Gly Ser Leu Asn Glu
      95      100      105
Leu Leu His Arg Lys Thr Glu Tyr Pro Asp Val Ala Trp Pro Leu
      110      115      120
Arg Phe Arg Ile Leu His Glu Ile Ala Leu Gly Val Asn Tyr Leu

```

	125		130		135
His Asn Met Thr	Pro Pro Leu Leu His	His Asp Leu Lys Thr	Gln		
	140		145		150
Asn Ile Leu Leu	Asp Asn Glu Phe His	Val Lys Ile Ala Asp	Phe		
	155		160		165
Gly Leu Ser Lys	Trp Arg Met Met Ser	Leu Ser Gln Ser Arg	Ser		
	170		175		180
Ser Lys Ser Ala	Pro Glu Gly Gly Thr	Ile Ile Tyr Met Pro	Pro		
	185		190		195
Glu Asn Tyr Glu	Pro Gly Gln Lys Ser	Arg Ala Ser Ile Lys	His		
	200		205		210
Asp Ile Tyr Ser	Tyr Ala Val Ile Thr	Trp Glu Val Leu Ser	Arg		
	215		220		225
Lys Gln Pro Phe	Glu Asp Val Thr Asn	Pro Leu Gln Ile Met	Tyr		
	230		235		240
Ser Val Ser Gln	Gly His Arg Pro Val	Ile Asn Glu Glu Ser	Leu		
	245		250		255
Pro Tyr Asp Ile	Pro His Arg Ala Arg	Met Ile Ser Leu Ile	Glu		
	260		265		270
Ser Gly Trp Ala	Gln Asn Pro Asp Glu	Arg Pro Ser Phe Leu	Lys		
	275		280		285
Cys Leu Ile Glu	Leu Glu Pro Val Leu	Arg Thr Phe Glu Glu	Ile		
	290		295		300
Thr Phe Leu Glu	Ala Val Ile Gln Leu	Lys Lys Thr Lys Leu	Gln		
	305		310		315
Ser Val Ser Ser	Ala Ile His Leu Cys	Asp Lys Lys Lys Met	Glu		
	320		325		330
Leu Ser Leu Asn	Ile Pro Val Asn His	Gly Pro Gln Glu Glu	Ser		
	335		340		345
Cys Gly Ser Ser	Gln Leu His Glu Asn	Ser Gly Ser Pro Glu	Thr		
	350		355		360
Ser Arg Ser Leu	Pro Ala Pro Gln Asp	Asn Asp Phe Leu Ser	Arg		
	365		370		375
Lys Ala Gln Asp	Cys Tyr Phe Met Lys	Leu His His Cys Pro	Gly		
	380		385		390
Asn His Ser Trp	Asp Ser Thr Ile Ser	Gly Ser Gln Arg Ala	Ala		
	395		400		405
Phe Cys Asp His	Lys Thr Thr Pro Cys	Ser Ser Ala Ile Ile	Asn		
	410		415		420
Pro Leu Ser Thr	Ala Gly Asn Ser Glu	Arg Leu Gln Pro Gly	Ile		
	425		430		435
Ala Gln Gln Trp	Ile Gln Ser Lys Arg	Glu Asp Ile Val Asn	Gln		
	440		445		450
Met Thr Glu Ala	Cys Leu Asn Gln Ser	Leu Asp Ala Leu Leu	Ser		
	455		460		465
Arg Asp Leu Ile	Met Lys Glu Asp Tyr	Glu Leu Val Ser Thr	Lys		
	470		475		480
Pro Thr Arg Thr	Ser Lys Val Arg Gln	Leu Leu Asp Thr Thr	Asp		
	485		490		495
Ile Gln Gly Glu	Glu Phe Ala Lys Val	Ile Val Gln Lys Leu	Lys		
	500		505		510
Asp Asn Lys Gln	Met Gly Leu Gln Pro	Tyr Pro Glu Ile Leu	Val		
	515		520		525
Val Ser Arg Ser	Pro Ser Leu Asn Leu	Leu Gln Asn Lys Ser	Met		
	530		535		540

<210> 7  
 <211> 454  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone Number: 2883243

<400> 7  
 Met Tyr Asn Thr Val Trp Asn Met Glu Asp Leu Asp Leu Glu Tyr  
 1 5 10 15  
 Ala Lys Thr Asp Ile Asn Cys Gly Thr Asp Leu Met Phe Tyr Ile  
 20 25 30  
 Glu Met Asp Pro Pro Ala Leu Pro Pro Lys Pro Pro Lys Pro Thr  
 35 40 45  
 Thr Val Ala Asn Asn Gly Met Asn Asn Asn Met Ser Leu Gln Asp  
 50 55 60  
 Ala Glu Trp Tyr Trp Gly Asp Ile Ser Arg Glu Glu Val Asn Glu  
 65 70 75  
 Lys Leu Arg Asp Thr Ala Asp Gly Thr Phe Leu Val Arg Asp Ala  
 80 85 90  
 Ser Thr Lys Met His Gly Asp Tyr Thr Leu Thr Leu Arg Lys Gly  
 95 100 105  
 Gly Asn Asn Lys Leu Ile Lys Ile Phe His Arg Asp Gly Lys Tyr  
 110 115 120  
 Gly Phe Ser Asp Pro Leu Thr Phe Ser Ser Val Val Glu Leu Ile  
 125 130 135  
 Asn His Tyr Arg Asn Glu Ser Leu Ala Gln Tyr Asn Pro Lys Leu  
 140 145 150  
 Asp Val Lys Leu Leu Tyr Pro Val Ser Lys Tyr Gln Gln Asp Gln  
 155 160 165  
 Val Val Lys Glu Asp Asn Ile Glu Ala Val Gly Lys Lys Leu His  
 170 175 180  
 Glu Tyr Asn Thr Gln Phe Gln Glu Lys Ser Arg Glu Tyr Asp Arg  
 185 190 195  
 Leu Tyr Glu Glu Tyr Thr Arg Thr Ser Gln Glu Ile Gln Met Lys  
 200 205 210  
 Arg Thr Ala Ile Glu Ala Phe Asn Glu Thr Ile Lys Ile Phe Glu  
 215 220 225  
 Glu Gln Cys Gln Thr Gln Glu Arg Tyr Ser Lys Glu Tyr Ile Glu  
 230 235 240  
 Lys Phe Lys Arg Glu Gly Asn Glu Lys Glu Ile Gln Arg Ile Met  
 245 250 255  
 His Asn Tyr Asp Lys Leu Lys Ser Arg Ile Ser Glu Ile Ile Asp  
 260 265 270  
 Ser Arg Arg Arg Leu Glu Glu Asp Leu Lys Lys Gln Ala Ala Glu  
 275 280 285  
 Tyr Arg Glu Ile Asp Lys Arg Met Asn Ser Ile Lys Pro Asp Leu  
 290 295 300  
 Ile Gln Leu Arg Lys Thr Arg Asp Gln Tyr Leu Met Trp Leu Thr  
 305 310 315  
 Gln Lys Gly Val Arg Gln Lys Lys Leu Asn Glu Trp Leu Gly Asn  
 320 325 330  
 Glu Asn Thr Glu Asp Gln Tyr Ser Leu Val Glu Asp Asp Glu Asp  
 335 340 345

```

Leu Pro His His Asp Glu Lys Thr Trp Asn Val Gly Ser Ser Asn
      350                      355                      360
Arg Asn Lys Ala Glu Asn Leu Leu Arg Gly Lys Arg Asp Gly Thr
      365                      370                      375
Phe Leu Val Arg Glu Ser Ser Lys Gln Gly Cys Tyr Ala Cys Ser
      380                      385                      390
Val Val Val Asp Gly Glu Val Lys His Cys Val Ile Asn Lys Thr
      395                      400                      405
Ala Thr Gly Tyr Gly Phe Ala Glu Pro Tyr Asn Leu Tyr Ser Ser
      410                      415                      420
Leu Lys Glu Leu Val Leu His Tyr Gln His Thr Ser Leu Val Gln
      425                      430                      435
His Asn Asp Ser Leu Asn Val Thr Leu Ala Tyr Pro Val Tyr Ala
      440                      445                      450
Gln Gln Arg Arg

```

&lt;210&gt; 8

&lt;211&gt; 502

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 3173355

&lt;400&gt; 8

```

Met Phe Gly Thr Leu Leu Leu Tyr Cys Phe Phe Leu Ala Thr Val
  1                      5                      10                      15
Pro Ala Leu Ala Glu Thr Gly Gly Glu Arg Gln Leu Ser Pro Glu
      20                      25                      30
Lys Ser Glu Ile Trp Gly Pro Gly Leu Lys Ala Asp Val Val Leu
      35                      40                      45
Pro Ala Arg Tyr Phe Tyr Ile Gln Ala Val Asp Thr Ser Gly Asn
      50                      55                      60
Lys Phe Thr Ser Ser Pro Gly Glu Lys Val Phe Gln Val Lys Val
      65                      70                      75
Ser Ala Pro Glu Glu Gln Phe Thr Arg Val Gly Val Gln Val Leu
      80                      85                      90
Asp Arg Lys Asp Gly Ser Phe Ile Val Arg Tyr Arg Met Tyr Ala
      95                      100                     105
Ser Tyr Lys Asn Leu Lys Val Glu Ile Lys Phe Gln Gly Gln His
      110                     115                     120
Val Ala Lys Ser Pro Tyr Ile Leu Lys Gly Pro Val Tyr His Glu
      125                     130                     135
Asn Cys Asp Cys Pro Leu Gln Asp Ser Ala Ala Trp Leu Arg Glu
      140                     145                     150
Met Asn Cys Pro Glu Thr Ile Ala Gln Ile Gln Arg Asp Leu Ala
      155                     160                     165
His Phe Pro Ala Val Asp Pro Glu Lys Ile Ala Val Glu Ile Pro
      170                     175                     180
Lys Arg Phe Gly Gln Arg Gln Ser Leu Cys His Tyr Thr Leu Lys
      185                     190                     195
Asp Asn Lys Val Tyr Ile Lys Thr His Gly Glu His Val Gly Phe
      200                     205                     210
Arg Ile Phe Met Asp Ala Ile Leu Leu Ser Leu Thr Arg Lys Val

```

215	220	225
Lys Met Pro Asp Val Glu Leu Phe Val	Asn Leu Gly Asp Trp Pro	
230	235	240
Leu Glu Lys Lys Lys Ser Asn Ser Asn	Ile His Pro Ile Phe Ser	
245	250	255
Trp Cys Gly Ser Thr Asp Ser Lys Asp	Ile Val Met Pro Thr Tyr	
260	265	270
Asp Leu Thr Asp Ser Val Leu Glu Thr	Met Gly Arg Val Ser Leu	
275	280	285
Asp Met Met Ser Val Gln Ala Asn Thr	Gly Pro Pro Trp Glu Ser	
290	295	300
Lys Asn Ser Thr Ala Val Trp Arg Gly	Arg Asp Ser Arg Lys Glu	
305	310	315
Arg Leu Glu Leu Val Lys Leu Ser Arg	Lys His Pro Glu Leu Ile	
320	325	330
Asp Ala Ala Phe Thr Asn Phe Phe Phe	Phe Lys His Asp Glu Asn	
335	340	345
Leu Tyr Gly Pro Ile Val Lys His Ile	Ser Phe Phe Asp Phe Phe	
350	355	360
Lys His Lys Tyr Gln Ile Asn Ile Asp	Gly Thr Val Ala Ala Tyr	
365	370	375
Arg Leu Pro Tyr Leu Leu Val Gly Asp	Ser Val Val Leu Lys Gln	
380	385	390
Asp Ser Ile Tyr Tyr Glu His Phe Tyr	Asn Glu Leu Gln Pro Trp	
395	400	405
Lys His Tyr Ile Pro Val Lys Ser Asn	Leu Ser Asp Leu Leu Glu	
410	415	420
Lys Leu Lys Trp Ala Lys Asp His Asp	Glu Glu Ala Lys Lys Ile	
425	430	435
Ala Lys Ala Gly Gln Glu Phe Ala Arg	Asn Asn Leu Met Gly Asp	
440	445	450
Asp Ile Phe Cys Tyr Tyr Phe Lys Leu	Phe Gln Glu Tyr Ala Asn	
455	460	465
Leu Gln Val Ser Glu Pro Gln Ile Arg	Glu Gly Met Lys Arg Val	
470	475	480
Glu Pro Gln Thr Glu Asp Asp Leu Phe	Pro Cys Thr Cys His Arg	
485	490	495
Lys Lys Thr Lys Asp Glu Leu		
500		

&lt;210&gt; 9

&lt;211&gt; 282

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 5116906

&lt;400&gt; 9

Met Trp Ala Cys Gly Val Ile Leu Tyr	Ile Leu Leu Val Gly Tyr
1	5 10 15
Pro Pro Phe Trp Asp Glu Asp Gln His	Arg Leu Tyr Gln Gln Ile
20	25 30
Lys Ala Gly Ala Tyr Asp Phe Pro Ser	Pro Glu Trp Asp Thr Val
35	40 45



```

Thr Pro Glu Ala Lys Asp Leu Ile Asn Lys Met Leu Thr Ile Asn
      50                      55                      60
Pro Ala Lys Arg Ile Thr Ala Ser Glu Ala Leu Lys His Pro Trp
      65                      70                      75
Ile Cys Gln Arg Ser Thr Val Ala Ser Met Met His Arg Gln Glu
      80                      85                      90
Thr Val Asp Cys Leu Lys Lys Phe Asn Ala Arg Arg Lys Leu Lys
      95                      100                     105
Gly Ala Ile Leu Thr Thr Met Leu Ala Thr Arg Asn Phe Ser Ala
      110                     115                     120
Ala Lys Ser Leu Leu Lys Lys Pro Asp Gly Val Lys Glu Ser Thr
      125                     130                     135
Glu Ser Ser Asn Thr Thr Ile Glu Asp Glu Asp Val Lys Ala Arg
      140                     145                     150
Lys Gln Glu Ile Ile Lys Val Thr Glu Gln Leu Ile Glu Ala Ile
      155                     160                     165
Asn Asn Gly Asp Phe Glu Ala Tyr Thr Lys Ile Cys Asp Pro Gly
      170                     175                     180
Leu Thr Ala Phe Glu Pro Glu Ala Leu Gly Asn Leu Val Glu Gly
      185                     190                     195
Met Asp Phe His Arg Phe Tyr Phe Glu Asn Ala Leu Ser Lys Ser
      200                     205                     210
Asn Lys Pro Ile His Thr Ile Ile Leu Asn Pro His Val His Leu
      215                     220                     225
Val Gly Asp Asp Ala Ala Cys Ile Ala Tyr Ile Arg Leu Thr Gln
      230                     235                     240
Tyr Met Asp Gly Ser Gly Met Pro Lys Thr Met Gln Ser Glu Glu
      245                     250                     255
Thr Arg Val Trp His Arg Arg Asp Gly Lys Trp Gln Asn Val His
      260                     265                     270
Phe His Arg Ser Gly Ser Pro Thr Val Pro Ile Asn
      275                     280

```

&lt;210&gt; 10

&lt;211&gt; 510

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 940589

&lt;400&gt; 10

```

Met Lys Ala Asp Ile Lys Ile Trp Ile Leu Thr Gly Asp Lys Gln
      1           5           10           15
Glu Thr Ala Ile Asn Ile Gly His Ser Cys Lys Leu Leu Lys Lys
      20           25           30
Asn Met Gly Met Ile Val Ile Asn Glu Gly Ser Leu Asp Ser Phe
      35           40           45
Ser Asn Thr Gln Asn Ser Arg Lys Glu Ala Val Leu Leu Ala Lys
      50           55           60
Met Lys His Pro Asn Ile Val Ala Phe Lys Glu Ser Phe Glu Ala
      65           70           75
Glu Gly His Leu Tyr Ile Val Met Glu Tyr Cys Asp Gly Gly Asp
      80           85           90

```

Leu Met Gln Lys Ile Lys Gln Gln Lys Gly Lys Leu Phe Pro Glu	95	100	105
Asp Met Ile Leu Asn Trp Phe Thr Gln Met Cys Leu Gly Val Asn	110	115	120
His Ile His Lys Lys Arg Val Leu His Arg Asp Ile Lys Ser Lys	125	130	135
Asn Ile Phe Leu Thr Gln Asn Gly Lys Val Lys Leu Gly Asp Phe	140	145	150
Gly Ser Ala Arg Leu Leu Ser Asn Pro Met Ala Phe Ala Cys Thr	155	160	165
Tyr Val Gly Thr Pro Tyr Tyr Val Pro Pro Glu Ile Trp Glu Asn	170	175	180
Leu Pro Tyr Asn Asn Lys Ser Asp Ile Trp Ser Leu Gly Cys Ile	185	190	195
Leu Tyr Glu Leu Cys Thr Leu Lys His Pro Phe Gln Ala Asn Ser	200	205	210
Trp Lys Asn Leu Ile Leu Lys Val Cys Gln Gly Cys Ile Ser Pro	215	220	225
Leu Pro Ser His Tyr Ser Tyr Glu Leu Gln Phe Leu Val Lys Gln	230	235	240
Met Phe Lys Arg Asn Pro Ser His Arg Pro Ser Ala Thr Thr Leu	245	250	255
Leu Ser Arg Gly Ile Val Ala Arg Leu Val Gln Lys Cys Leu Pro	260	265	270
Pro Glu Ile Ile Met Glu Tyr Gly Glu Glu Val Leu Glu Glu Ile	275	280	285
Lys Asn Ser Lys His Asn Thr Pro Arg Lys Lys Thr Asn Pro Ser	290	295	300
Arg Ile Arg Ile Ala Leu Gly Asn Glu Ala Ser Thr Val Gln Glu	305	310	315
Glu Glu Gln Asp Arg Lys Gly Ser His Thr Asp Leu Glu Ser Ile	320	325	330
Asn Glu Asn Leu Val Glu Ser Ala Leu Arg Arg Val Asn Arg Glu	335	340	345
Glu Lys Gly Asn Lys Ser Val His Leu Arg Lys Ala Ser Ser Pro	350	355	360
Asn Leu His Arg Arg Gln Trp Glu Lys Asn Val Pro Asn Thr Ala	365	370	375
Leu Thr Ala Leu Glu Asn Ala Ser Ile Leu Thr Ser Ser Leu Thr	380	385	390
Ala Glu Asp Asp Arg Gly Gly Ser Val Ile Lys Tyr Ser Lys Asn	395	400	405
Thr Thr Arg Lys Gln Trp Leu Lys Glu Thr Pro Asp Thr Leu Leu	410	415	420
Asn Ile Leu Lys Asn Ala Asp Leu Ser Leu Ala Phe Gln Thr Tyr	425	430	435
Thr Ile Tyr Arg Pro Gly Ser Glu Gly Phe Leu Lys Gly Pro Leu	440	445	450
Ser Glu Glu Thr Glu Ala Ser Asp Ser Val Asp Gly Gly His Asp	455	460	465
Ser Val Ile Leu Asp Pro Glu Arg Leu Glu Pro Gly Leu Asp Glu	470	475	480
Glu Asp Thr Asp Phe Glu Glu Glu Asp Asp Asn Pro Asp Trp Val	485	490	495
Ser Glu Leu Lys Lys Arg Ala Gly Trp Gln Gly Leu Cys Asp Arg	500	505	510

<210> 11  
 <211> 248  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone Number: 304421

<400> 11  
 Met Ala Glu Thr Ser Leu Pro Glu Leu Gly Gly Glu Asp Lys Ala  
 1 5 10 15  
 Thr Pro Cys Pro Ser Ile Leu Glu Leu Glu Glu Leu Leu Arg Ala  
 20 25 30  
 Gly Lys Ser Ser Cys Ser Arg Val Asp Glu Val Trp Pro Asn Leu  
 35 40 45  
 Phe Ile Gly Asp Ala Met Asp Ser Leu Gln Lys Gln Asp Leu Arg  
 50 55 60  
 Arg Pro Lys Ile His Gly Ala Val Gln Ala Ser Pro Tyr Gln Pro  
 65 70 75  
 Pro Thr Leu Ala Ser Leu Gln Arg Leu Leu Trp Val Arg Gln Ala  
 80 85 90  
 Ala Thr Leu Asn His Ile Asp Glu Val Trp Pro Ser Leu Phe Leu  
 95 100 105  
 Gly Asp Ala Tyr Ala Ala Arg Asp Lys Ser Lys Leu Ile Gln Leu  
 110 115 120  
 Gly Ile Thr His Val Val Asn Ala Ala Ala Gly Lys Phe Gln Val  
 125 130 135  
 Asp Thr Gly Ala Lys Phe Tyr Arg Gly Met Ser Leu Glu Tyr Tyr  
 140 145 150  
 Gly Ile Glu Ala Asp Asp Asn Pro Phe Phe Asp Leu Ser Val Tyr  
 155 160 165  
 Phe Leu Pro Val Ala Arg Tyr Ile Arg Ala Ala Leu Ser Val Pro  
 170 175 180  
 Gln Gly Arg Val Leu Val His Cys Ala Met Gly Val Ser Arg Ser  
 185 190 195  
 Ala Thr Leu Val Leu Ala Phe Leu Met Ile Tyr Glu Asn Met Thr  
 200 205 210  
 Leu Val Glu Ala Ile Gln Thr Val Gln Ala His Arg Asn Ile Cys  
 215 220 225  
 Pro Asn Ser Gly Phe Leu Arg Gln Leu Gln Val Leu Asp Asn Arg  
 230 235 240  
 Leu Gly Arg Glu Thr Gly Arg Phe  
 245

<210> 12  
 <211> 810  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone Number: 1213802

&lt;400&gt; 12

Met	Pro	Asn	Gln	Gly	Glu	Asp	Cys	Tyr	Phe	Phe	Phe	Tyr	Ser	Thr
1				5					10					15
Cys	Thr	Lys	Gly	Asp	Ser	Cys	Pro	Phe	Arg	His	Cys	Glu	Ala	Ala
				20					25					30
Ile	Gly	Asn	Glu	Thr	Val	Cys	Thr	Leu	Trp	Gln	Glu	Gly	Arg	Cys
				35					40					45
Phe	Arg	Gln	Val	Cys	Arg	Phe	Arg	His	Met	Glu	Ile	Asp	Lys	Lys
				50					55					60
Arg	Ser	Glu	Ile	Pro	Cys	Tyr	Trp	Glu	Asn	Gln	Pro	Thr	Gly	Cys
				65					70					75
Gln	Lys	Leu	Asn	Cys	Ala	Phe	His	His	Asn	Arg	Gly	Arg	Tyr	Val
				80					85					90
Asp	Gly	Leu	Phe	Leu	Pro	Pro	Ser	Lys	Thr	Val	Leu	Pro	Thr	Val
				95					100					105
Pro	Glu	Ser	Pro	Glu	Glu	Glu	Val	Lys	Ala	Ser	Gln	Leu	Ser	Val
				110					115					120
Gln	Gln	Asn	Lys	Leu	Ser	Val	Gln	Ser	Asn	Pro	Ser	Pro	Gln	Leu
				125					130					135
Arg	Ser	Val	Met	Lys	Val	Glu	Ser	Ser	Glu	Asn	Val	Pro	Ser	Pro
				140					145					150
Thr	His	Pro	Pro	Val	Val	Ile	Asn	Ala	Ala	Asp	Asp	Asp	Glu	Asp
				155					160					165
Asp	Asp	Asp	Gln	Phe	Ser	Glu	Glu	Gly	Asp	Glu	Thr	Lys	Thr	Pro
				170					175					180
Thr	Leu	Gln	Pro	Thr	Pro	Glu	Val	His	Asn	Gly	Leu	Arg	Val	Thr
				185					190					195
Ser	Val	Arg	Lys	Pro	Ala	Val	Asn	Ile	Lys	Gln	Gly	Glu	Cys	Leu
				200					205					210
Asn	Phe	Gly	Ile	Lys	Thr	Leu	Glu	Glu	Ile	Lys	Ser	Lys	Lys	Met
				215					220					225
Lys	Glu	Lys	Ser	Lys	Lys	Gln	Gly	Glu	Gly	Ser	Ser	Gly	Val	Ser
				230					235					240
Ser	Leu	Leu	Leu	His	Pro	Glu	Pro	Val	Pro	Gly	Pro	Glu	Lys	Glu
				245					250					255
Asn	Val	Arg	Thr	Val	Val	Arg	Thr	Val	Thr	Leu	Ser	Thr	Lys	Gln
				260					265					270
Gly	Glu	Glu	Pro	Leu	Val	Arg	Leu	Ser	Leu	Thr	Glu	Arg	Leu	Gly
				275					280					285
Lys	Arg	Lys	Phe	Ser	Ala	Gly	Gly	Asp	Ser	Asp	Pro	Pro	Leu	Lys
				290					295					300
Arg	Ser	Leu	Ala	Gln	Arg	Leu	Gly	Lys	Lys	Val	Glu	Ala	Pro	Glu
				305					310					315
Thr	Asn	Ile	Asp	Lys	Thr	Pro	Lys	Lys	Ala	Gln	Val	Ser	Lys	Ser
				320					325					330
Leu	Lys	Glu	Arg	Leu	Gly	Met	Ser	Ala	Asp	Pro	Asp	Asn	Glu	Asp
				335					340					345
Ala	Thr	Asp	Lys	Val	Asn	Lys	Val	Gly	Glu	Ile	His	Val	Lys	Thr
				350					355					360
Leu	Glu	Glu	Ile	Leu	Leu	Glu	Arg	Ala	Ser	Gln	Lys	Arg	Gly	Glu
				365					370					375
Leu	Gln	Thr	Lys	Leu	Lys	Thr	Glu	Gly	Pro	Ser	Lys	Thr	Asp	Asp
				380					385					390
Ser	Thr	Ser	Gly	Ala	Arg	Ser	Ser	Ser	Thr	Ile	Arg	Ile	Lys	Thr
				395					400					405
Phe	Ser	Glu	Val	Leu	Ala	Glu	Lys	Lys	His	Arg	Gln	Gln	Glu	Ala

	410		415		420
Glu Arg Gln Lys	Ser Lys Lys Asp Thr	Thr Cys Ile Lys Leu Lys			
	425		430		435
Ile Asp Ser Glu	Ile Lys Lys Thr Val	Val Leu Pro Pro Ile Val			
	440		445		450
Ala Ser Arg Gly	Gln Ser Glu Glu Pro	Ala Gly Lys Thr Lys Ser			
	455		460		465
Met Gln Glu Val	His Ile Lys Thr Leu	Glu Glu Ile Lys Leu Glu			
	470		475		480
Lys Ala Leu Arg	Val Gln Gln Ser Ser	Glu Ser Ser Thr Ser Ser			
	485		490		495
Pro Ser Gln His	Glu Ala Thr Pro Gly	Ala Arg Arg Leu Leu Arg			
	500		505		510
Ile Thr Lys Arg	Thr Gly Met Lys Glu	Glu Lys Asn Leu Gln Glu			
	515		520		525
Gly Asn Glu Val	Asp Ser Gln Ser Ser	Ile Arg Thr Glu Ala Lys			
	530		535		540
Glu Ala Ser Gly	Glu Thr Thr Gly Val	Asp Ile Thr Lys Ile Gln			
	545		550		555
Val Lys Arg Cys	Glu Thr Met Arg Glu	Lys His Met Gln Lys Gln			
	560		565		570
Gln Glu Arg Glu	Lys Ser Val Leu Thr	Pro Leu Arg Gly Asp Val			
	575		580		585
Ala Ser Cys Asn	Thr Gln Val Ala Glu	Lys Pro Val Leu Thr Ala			
	590		595		600
Val Pro Gly Ile	Thr Arg His Leu Thr	Lys Arg Leu Pro Thr Lys			
	605		610		615
Ser Ser Gln Lys	Val Glu Val Glu Thr	Ser Gly Ile Gly Asp Ser			
	620		625		630
Leu Leu Asn Val	Lys Cys Ala Ala Gln	Thr Leu Glu Lys Arg Gly			
	635		640		645
Lys Ala Lys Pro	Lys Val Asn Val Lys	Pro Ser Val Val Lys Val			
	650		655		660
Val Ser Ser Pro	Lys Leu Ala Pro Lys	Arg Lys Ala Val Glu Met			
	665		670		675
His Ala Ala Val	Ile Ala Ala Val Lys	Pro Leu Ser Ser Ser Ser			
	680		685		690
Val Leu Gln Glu	Pro Pro Ala Lys Lys	Ala Ala Val Ala Val Val			
	695		700		705
Pro Leu Val Ser	Glu Asp Lys Ser Val	Thr Val Pro Glu Ala Glu			
	710		715		720
Asn Pro Arg Asp	Ser Leu Val Leu Pro	Pro Thr Gln Ser Ser Ser			
	725		730		735
Asp Ser Ser Pro	Pro Glu Val Ser Gly	Pro Ser Ser Ser Gln Met			
	740		745		750
Ser Met Lys Thr	Arg Arg Leu Ser Ser	Ala Ser Thr Gly Lys Pro			
	755		760		765
Pro Leu Ser Val	Glu Asp Asp Phe Glu	Lys Leu Ile Trp Glu Ile			
	770		775		780
Ser Gly Gly Lys	Leu Glu Ala Glu Ile	Asp Leu Asp Pro Gly Lys			
	785		790		795
Asp Glu Asp Asp	Leu Leu Leu Glu Leu	Ser Glu Met Ile Asp Ser			
	800		805		810

&lt;210&gt; 13

&lt;211&gt; 549

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 1378134

&lt;400&gt; 13

```

Met Arg Arg Arg Ala Ser Asn Ala Ala Ala Ala Ala His Thr Ile
 1          5          10          15
Gly Gly Ser Lys His Thr Met Asn Asp His Leu His Val Gly Ser
 20          25          30
His Ala His Gly Gln Ile Gln Val Arg Gln Leu Phe Glu Asp Asn
 35          40          45
Ser Asn Lys Arg Thr Val Leu Thr Thr Gln Pro Asn Gly Leu Thr
 50          55          60
Thr Val Gly Lys Thr Gly Leu Pro Val Val Pro Glu Arg Gln Leu
 65          70          75
Asp Ser Ile His Arg Arg Gln Gly Ser Ser Thr Ser Leu Lys Ser
 80          85          90
Met Glu Gly Met Gly Lys Val Lys Ala Thr Pro Met Thr Pro Glu
 95          100         105
Gln Ala Met Lys Gln Tyr Met Gln Lys Leu Thr Ala Phe Glu His
110          115         120
His Glu Ile Phe Ser Tyr Pro Glu Ile Tyr Phe Leu Gly Leu Asn
125          130         135
Ala Lys Lys Arg Gln Gly Met Thr Gly Gly Pro Asn Asn Gly Gly
140          145         150
Tyr Asp Asp Asp Gln Gly Ser Tyr Val Gln Val Pro His Asp His
155          160         165
Val Ala Tyr Arg Tyr Glu Val Leu Lys Val Ile Gly Lys Gly Ser
170          175         180
Phe Gly Gln Val Val Lys Ala Tyr Asp His Lys Val His Gln His
185          190         195
Val Ala Leu Lys Met Val Arg Asn Glu Lys Arg Phe His Arg Gln
200          205         210
Ala Ala Glu Glu Ile Arg Ile Leu Glu His Leu Arg Lys Gln Asp
215          220         225
Lys Asp Asn Thr Met Asn Val Ile His Met Leu Glu Asn Phe Thr
230          235         240
Phe Arg Asn His Ile Cys Met Thr Phe Glu Leu Leu Ser Met Asn
245          250         255
Leu Tyr Glu Leu Ile Lys Lys Asn Lys Phe Gln Gly Phe Ser Leu
260          265         270
Pro Leu Val Arg Lys Phe Ala His Ser Ile Leu Gln Cys Leu Asp
275          280         285
Ala Leu His Lys Asn Arg Ile Ile His Cys Asp Leu Lys Pro Glu
290          295         300
Asn Ile Leu Leu Lys Gln Gln Gly Arg Ser Gly Ile Lys Val Ile
305          310         315
Asp Phe Gly Ser Ser Cys Tyr Glu His Gln Arg Val Tyr Thr Tyr
320          325         330
Ile Gln Ser Arg Phe Tyr Arg Ala Pro Glu Val Ile Leu Gly Ala
335          340         345
Arg Tyr Gly Met Pro Ile Asp Met Trp Ser Leu Gly Cys Ile Leu

```

350	355	360
Ala Glu Leu Leu Thr Gly Tyr Pro Leu Leu Pro Gly Glu Asp Glu		
365	370	375
Gly Asp Gln Leu Ala Cys Met Ile Glu Leu Leu Gly Met Pro Ser		
380	385	390
Gln Lys Leu Leu Asp Ala Ser Lys Arg Ala Lys Asn Phe Val Ser		
395	400	405
Ser Lys Gly Tyr Pro Arg Tyr Cys Thr Val Thr Thr Leu Ser Asp		
410	415	420
Gly Ser Val Val Leu Asn Gly Gly Arg Ser Arg Arg Gly Lys Leu		
425	430	435
Arg Gly Pro Pro Glu Ser Arg Glu Trp Gly Asn Ala Leu Lys Gly		
440	445	450
Cys Asp Asp Pro Leu Phe Leu Asp Phe Leu Lys Gln Cys Leu Glu		
455	460	465
Trp Asp Pro Ala Val Arg Met Thr Pro Gly Gln Ala Leu Arg His		
470	475	480
Pro Trp Leu Arg Arg Arg Leu Pro Lys Pro Pro Thr Gly Glu Lys		
485	490	495
Thr Ser Val Lys Arg Ile Thr Glu Ser Thr Gly Ala Ile Thr Ser		
500	505	510
Ile Ser Lys Leu Pro Pro Pro Ser Ser Ser Ala Ser Lys Leu Arg		
515	520	525
Thr Asn Leu Ala Gln Met Thr Asp Ala Asn Gly Asn Ile Gln Gln		
530	535	540
Arg Thr Val Leu Pro Lys Leu Val Ser		
545		

&lt;210&gt; 14

&lt;211&gt; 416

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 1490070

&lt;400&gt; 14

Met Met Pro Gln Leu Gln Phe Lys Asp Ala Phe Trp Cys Arg Asp		
1	5	10
Phe Thr Ala His Thr Gly Tyr Glu Val Leu Leu Gln Arg Leu Leu		
20	25	30
Asp Gly Arg Lys Met Cys Lys Asp Met Val Glu Leu Leu Trp Gln		
35	40	45
Arg Ala Gln Ala Glu Glu Arg Tyr Gly Lys Glu Leu Val Gln Ile		
50	55	60
Ala Arg Lys Ala Gly Gly Gln Thr Glu Ile Asn Ser Leu Arg Ala		
65	70	75
Ser Phe Asp Ser Leu Lys Gln Gln Met Glu Asn Val Gly Ser Ser		
80	85	90
His Ile Gln Leu Ala Leu Thr Leu Arg Glu Glu Leu Arg Ser Leu		
95	100	105
Glu Glu Phe Arg Glu Arg Gln Lys Glu Gln Arg Lys Lys Tyr Glu		
110	115	120
Ala Val Met Asp Arg Val Gln Lys Ser Lys Leu Ser Leu Tyr Lys		

	125		130		135
Lys Ala Met Glu Ser Lys Lys Thr Tyr Glu Gln Lys Cys Arg Asp					
	140		145		150
Ala Asp Asp Ala Glu Gln Ala Phe Glu Arg Ile Ser Ala Asn Gly					
	155		160		165
His Gln Lys Gln Val Glu Lys Ser Gln Asn Lys Ala Arg Gln Cys					
	170		175		180
Lys Asp Ser Ala Thr Glu Ala Glu Arg Val Tyr Arg Gln Ser Ile					
	185		190		195
Ala Gln Leu Glu Lys Val Arg Ala Glu Trp Glu Gln Glu His Arg					
	200		205		210
Thr Thr Cys Glu Ala Phe Gln Leu Gln Glu Phe Asp Arg Leu Thr					
	215		220		225
Ile Leu Arg Asn Ala Leu Trp Val His Ser Asn Gln Leu Ser Met					
	230		235		240
Gln Cys Val Lys Asp Asp Glu Leu Tyr Glu Glu Val Arg Leu Thr					
	245		250		255
Leu Glu Gly Cys Ser Ile Asp Ala Asp Ile Asp Ser Phe Ile Gln					
	260		265		270
Ala Lys Ser Thr Gly Thr Glu Pro Pro Ala Pro Val Pro Tyr Gln					
	275		280		285
Asn Tyr Tyr Asp Arg Glu Val Thr Pro Leu Thr Ser Ser Pro Gly					
	290		295		300
Ile Gln Pro Ser Cys Gly Met Ile Lys Arg Phe Ser Gly Leu Leu					
	305		310		315
His Gly Ser Pro Lys Thr Thr Ser Leu Ala Ala Ser Ala Ala Ser					
	320		325		330
Thr Glu Thr Leu Thr Pro Thr Pro Glu Arg Asn Glu Gly Val Tyr					
	335		340		345
Thr Ala Ile Ala Val Gln Glu Ile Gln Gly Asn Pro Ala Ser Pro					
	350		355		360
Ala Gln Glu Tyr Arg Ala Leu Tyr Asp Tyr Thr Ala Gln Asn Pro					
	365		370		375
Asp Glu Leu Asp Leu Ser Ala Gly Asp Ile Leu Glu Val Ile Leu					
	380		385		390
Glu Gly Glu Asp Gly Trp Trp Thr Val Glu Arg Asn Gly Gln Arg					
	395		400		405
Gly Phe Val Pro Gly Ser Tyr Leu Glu Lys Leu					
	410		415		

&lt;210&gt; 15

&lt;211&gt; 425

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 1997814

&lt;400&gt; 15

Met Glu Gln Gly Leu Glu Glu Glu Glu Val Asp Pro Arg Ile			
1	5	10	15
Gln Gly Glu Leu Glu Lys Leu Asn Gln Ser Thr Asp Asp Ile Asn			
	20	25	30
Arg Arg Glu Thr Glu Leu Glu Asp Ala Arg Gln Lys Phe Arg Ser			
	35	40	45



Val	Leu	Val	Glu	Ala	Thr	Val	Lys	Leu	Asp	Glu	Leu	Val	Lys	Lys	50	55	60
Ile	Gly	Lys	Ala	Val	Glu	Asp	Ser	Lys	Pro	Tyr	Trp	Glu	Ala	Arg	65	70	75
Arg	Val	Ala	Arg	Gln	Ala	Gln	Leu	Glu	Ala	Gln	Lys	Ala	Thr	Gln	80	85	90
Asp	Phe	Gln	Arg	Ala	Thr	Glu	Val	Leu	Arg	Ala	Ala	Lys	Glu	Thr	95	100	105
Ile	Ser	Leu	Ala	Glu	Gln	Arg	Leu	Leu	Glu	Asp	Asp	Lys	Arg	Gln	110	115	120
Phe	Asp	Ser	Ala	Trp	Gln	Glu	Met	Leu	Asn	His	Ala	Thr	Gln	Arg	125	130	135
Val	Met	Glu	Ala	Glu	Gln	Thr	Lys	Thr	Arg	Ser	Glu	Leu	Val	His	140	145	150
Lys	Glu	Thr	Ala	Ala	Arg	Tyr	Asn	Ala	Ala	Met	Gly	Arg	Met	Arg	155	160	165
Gln	Leu	Glu	Lys	Lys	Leu	Lys	Arg	Ala	Ile	Asn	Lys	Ser	Lys	Pro	170	175	180
Tyr	Phe	Glu	Leu	Lys	Ala	Lys	Tyr	Tyr	Val	Gln	Leu	Glu	Gln	Leu	185	190	195
Lys	Lys	Thr	Val	Asp	Asp	Leu	Gln	Ala	Lys	Leu	Thr	Leu	Ala	Lys	200	205	210
Gly	Glu	Tyr	Lys	Met	Ala	Leu	Lys	Asn	Leu	Glu	Met	Ile	Ser	Asp	215	220	225
Glu	Ile	His	Glu	Arg	Arg	Arg	Ser	Ser	Ala	Met	Gly	Pro	Arg	Gly	230	235	240
Cys	Gly	Val	Gly	Ala	Glu	Gly	Ser	Ser	Thr	Ser	Val	Glu	Asp	Leu	245	250	255
Pro	Gly	Ser	Lys	Pro	Glu	Pro	Asp	Ala	Ile	Ser	Val	Ala	Ser	Glu	260	265	270
Ala	Phe	Glu	Asp	Asp	Ser	Cys	Ser	Asn	Phe	Val	Ser	Glu	Asp	Asp	275	280	285
Ser	Glu	Thr	Gln	Ser	Val	Ser	Ser	Phe	Ser	Ser	Gly	Pro	Thr	Ser	290	295	300
Pro	Ser	Glu	Met	Pro	Asp	Gln	Phe	Pro	Ala	Val	Val	Arg	Pro	Gly	305	310	315
Ser	Leu	Asp	Leu	Pro	Ser	Pro	Val	Ser	Leu	Ser	Glu	Phe	Gly	Met	320	325	330
Met	Phe	Pro	Val	Leu	Gly	Pro	Arg	Ser	Glu	Cys	Ser	Gly	Ala	Ser	335	340	345
Ser	Pro	Glu	Cys	Glu	Val	Glu	Arg	Gly	Asp	Arg	Ala	Glu	Gly	Ala	350	355	360
Glu	Asn	Lys	Thr	Ser	Asp	Lys	Ala	Asn	Asn	Asn	Arg	Gly	Leu	Ser	365	370	375
Ser	Ser	Ser	Gly	Ser	Gly	Gly	Ser	Ser	Lys	Ser	Gln	Ser	Ser	Thr	380	385	390
Ser	Pro	Glu	Gly	Gln	Ala	Leu	Glu	Asn	Arg	Met	Lys	Gln	Leu	Ser	395	400	405
Leu	Gln	Cys	Ser	Lys	Gly	Arg	Asp	Gly	Ile	Ile	Ala	Asp	Ile	Lys	410	415	420
Met	Val	Gln	Ile	Gly											425		

&lt;210&gt; 16

&lt;211&gt; 1135

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2299715

&lt;400&gt; 16

```

Met Ala Asn Asp Ser Pro Ala Lys Ser Leu Val Asp Ile Asp Leu
  1          5          10          15
Ser Ser Leu Arg Asp Pro Ala Gly Ile Phe Glu Leu Val Glu Val
  20          25          30
Val Gly Asn Gly Thr Tyr Gly Gln Val Tyr Lys Gly Arg His Val
  35          40          45
Lys Thr Gly Gln Leu Ala Ala Ile Lys Val Met Asp Val Thr Glu
  50          55          60

Asp Glu Glu Glu Glu Ile Lys Leu Glu Ile Asn Met Leu Lys Lys
  65          70          75
Tyr Ser His His Arg Asn Ile Ala Thr Tyr Tyr Gly Ala Phe Ile
  80          85          90
Lys Lys Ser Pro Gly His Asp Asp Gln Leu Trp Leu Val Met
  95          100         105
Glu Phe Cys Gly Ala Gly Ser Ile Thr Asp Leu Val Lys Asn Thr
  110         115         120
Lys Gly Asn Thr Leu Lys Glu Asp Trp Ile Ala Tyr Ile Ser Arg
  125         130         135
Glu Ile Leu Arg Gly Leu Ala His Leu His Ile His His Val Ile
  140         145         150
His Arg Asp Ile Lys Gly Gln Asn Val Leu Leu Thr Glu Asn Ala
  155         160         165
Gly Val Lys Leu Val Asp Phe Gly Val Ser Ala Gln Leu Asp Arg
  170         175         180
Thr Val Gly Arg Arg Asn Thr Phe Ile Gly Thr Pro Tyr Trp Met
  185         190         195
Ala Pro Glu Val Ile Ala Cys Asp Glu Asn Pro Asp Ala Thr Tyr
  200         205         210
Asp Tyr Arg Ser Asp Leu Trp Ser Cys Gly Ile Thr Ala Ile Glu
  215         220         225
Met Ala Glu Gly Ala Pro Pro Leu Cys Asp Met His Pro Met Arg
  230         235         240
Ala Leu Phe Leu Ile Pro Arg Asn Pro Pro Pro Arg Leu Lys Ser
  245         250         255
Lys Lys Trp Ser Lys Lys Phe Phe Ser Phe Ile Glu Gly Cys Leu
  260         265         270
Val Lys Asn Tyr Met Gln Arg Pro Ser Thr Glu Gln Leu Leu Lys
  275         280         285
His Pro Phe Ile Arg Asp Gln Pro Asn Glu Arg Gln Val Arg Ile
  290         295         300
Gln Leu Lys Asp His Ile Asp Arg Thr Arg Lys Lys Arg Gly Glu
  305         310         315
Lys Asp Glu Thr Glu Tyr Glu Tyr Ser Gly Ser Glu Glu Glu Glu
  320         325         330
Glu Glu Val Pro Glu Gln Glu Gly Glu Pro Ser Ser Ile Val Asn
  335         340         345
Val Pro Gly Glu Ser Thr Leu Arg Arg Asp Phe Leu Arg Leu Gln

```

	350		355		360
Gln Glu Asn Lys	Glu Arg Ser Glu Ala	Leu Arg Arg Gln Gln	Leu		
	365		370		375
Leu Gln Glu Gln	Gln Leu Arg Glu Gln	Glu Glu Tyr Lys Arg	Gln		
	380		385		390
Leu Leu Ala Glu	Arg Gln Lys Arg Ile	Glu Gln Gln Lys Glu	Gln		
	395		400		405
Arg Arg Arg Leu	Glu Glu Gln Gln Arg	Arg Glu Arg Glu Ala	Arg		
	410		415		420
Arg Gln Gln Glu	Arg Glu Gln Arg Arg	Arg Glu Gln Glu Glu	Lys		
	425		430		435
Arg Arg Leu Glu	Glu Leu Glu Arg Arg	Arg Lys Glu Glu Glu	Glu		
	440		445		450
Arg Arg Arg Ala	Glu Glu Glu Lys Arg	Arg Val Glu Arg Glu	Gln		
	455		460		465
Glu Tyr Ile Arg	Arg Gln Leu Glu Glu	Glu Gln Arg His Leu	Glu		
	470		475		480
Val Leu Gln Gln	Gln Leu Leu Gln Glu	Gln Ala Met Leu Leu	His		
	485		490		495
Asp His Arg Arg	Pro His Pro Gln His	Ser Gln Gln Pro Pro	Pro		
	500		505		510
Pro Gln Gln Glu	Arg Ser Lys Pro Ser	Phe His Ala Pro Glu	Pro		
	515		520		525
Lys Ala His Tyr	Glu Pro Ala Asp Arg	Ala Arg Glu Val Pro	Val		
	530		535		540
Arg Thr Thr Ser	Arg Ser Pro Val Leu	Ser Arg Arg Asp Ser	Pro		
	545		550		555
Leu Gln Gly Ser	Gly Gln Gln Asn Ser	Gln Ala Gly Gln Arg	Asn		
	560		565		570
Ser Thr Ser Ile	Glu Pro Arg Leu Leu	Trp Glu Arg Val Glu	Lys		
	575		580		585
Leu Val Pro Arg	Pro Gly Ser Gly Ser	Ser Ser Gly Ser Ser	Asn		
	590		595		600
Ser Gly Ser Gln	Pro Gly Ser His Pro	Gly Ser Gln Ser Gly	Ser		
	605		610		615
Gly Glu Arg Phe	Arg Val Arg Ser Ser	Ser Lys Ser Glu Gly	Ser		
	620		625		630
Pro Ser Gln Arg	Leu Glu Asn Ala Val	Lys Lys Pro Glu Asp	Lys		
	635		640		645
Lys Glu Val Phe	Arg Pro Leu Lys Pro	Ala Asp Leu Thr Ala	Leu		
	650		655		660
Ala Lys Glu Leu	Arg Ala Val Glu Asp	Val Arg Pro Pro His	Lys		
	665		670		675
Val Thr Asp Tyr	Ser Ser Ser Ser Glu	Glu Ser Gly Thr Thr	Asp		
	680		685		690
Glu Glu Asp Asp	Asp Val Glu Gln Glu	Gly Ala Asp Glu Ser	Thr		
	695		700		705
Ser Gly Pro Glu	Asp Thr Arg Ala Ala	Ser Ser Leu Asn Leu	Ser		
	710		715		720
Asn Gly Glu Thr	Glu Ser Val Lys Thr	Met Ile Val His Asp	Asp		
	725		730		735
Val Glu Ser Glu	Pro Ala Met Thr Pro	Ser Lys Glu Gly Thr	Leu		
	740		745		750
Ile Val Arg Gln	Thr Gln Ser Ala Ser	Ser Thr Leu Gln Lys	His		
	755		760		765
Lys Ser Ser Ser	Ser Phe Thr Pro Phe	Ile Asp Pro Arg Leu	Leu		

	770		775		780
Gln Ile Ser Pro	Ser Ser Gly Thr Thr	Val Thr Ser Val Val	Gly		
	785		790		795
Phe Ser Cys Asp	Gly Met Arg Pro Glu Ala	Ile Arg Gln Asp Pro			
	800		805		810
Thr Arg Lys Gly	Ser Val Val Asn Val Asn	Pro Thr Asn Thr Arg			
	815		820		825
Pro Gln Ser Asp	Thr Pro Glu Ile Arg Lys	Tyr Lys Lys Arg Phe			
	830		835		840
Asn Ser Glu Ile	Leu Cys Ala Ala Leu Trp	Gly Val Asn Leu Leu			
	845		850		855
Val Gly Thr Glu	Ser Gly Leu Met Leu Leu	Asp Arg Ser Gly Gln			
	860		865		870
Gly Lys Val Tyr	Pro Leu Ile Asn Arg Arg	Arg Phe Gln Gln Met			
	875		880		885
Asp Val Leu Glu	Gly Leu Asn Val Leu Val	Thr Ile Ser Gly Lys			
	890		895		900
Lys Asp Lys Leu	Arg Val Tyr Tyr Leu Ser	Trp Leu Arg Asn Lys			
	905		910		915
Ile Leu His Asn	Asp Pro Glu Val Glu Lys	Lys Gln Gly Trp Thr			
	920		925		930
Thr Val Gly Asp	Leu Glu Gly Cys Val His	Tyr Lys Val Val Lys			
	935		940		945
Tyr Glu Arg Ile	Lys Phe Leu Val Ile Ala	Leu Lys Ser Ser Val			
	950		955		960
Glu Val Tyr Ala	Trp Ala Pro Lys Pro Tyr	His Lys Phe Met Ala			
	965		970		975
Phe Lys Ser Phe	Gly Glu Leu Val His Gly	Ser Cys Ala Gly Phe			
	980		985		990
His Ala Val Asp	Val Asp Ser Gly Ser Val	Tyr Asp Ile Tyr Leu			
	995		1000		1005
Pro Thr His Ile	Gln Cys Ser Ile Lys Pro	His Ala Ile Ile Ile			
	1010		1015		1020
Leu Pro Asn Thr	Asp Gly Met Glu Leu Leu	Val Cys Tyr Glu Asp			
	1025		1030		1035
Glu Gly Val Tyr	Val Asn Thr Tyr Gly Arg	Ile Thr Lys Asp Val			
	1040		1045		1050
Val Leu Gln Trp	Gly Glu Met Pro Thr Ser	Val Ala Tyr Ile Arg			
	1055		1060		1065
Ser Asn Gln Thr	Met Gly Trp Gly Glu Lys	Ala Ile Glu Ile Arg			
	1070		1075		1080
Ser Val Glu Thr	Gly His Leu Asp Gly Val	Phe Met His Lys Arg			
	1085		1090		1095
Ala Gln Arg Leu	Lys Phe Leu Cys Glu Arg	Asn Asp Lys Val Phe			
	1100		1105		1110
Phe Ala Ser Val	Arg Ser Gly Gly Ser Ser	Gln Val Tyr Phe Met			
	1115		1120		1125
Thr Leu Gly Arg	Thr Ser Leu Leu Ser Trp				
	1130		1135		

&lt;210&gt; 17

&lt;211&gt; 228

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone Number: 209854

<400> 17  
 Met Pro Thr Asn Cys Ala Ala Ala Gly Cys Ala Thr Thr Tyr Asn  
           1                  5                  10                  15  
 Lys His Ile Asn Ile Ser Phe His Arg Phe Pro Leu Asp Pro Lys  
                   20                  25                  30  
 Arg Arg Lys Glu Trp Val Arg Leu Val Arg Arg Lys Asn Phe Val  
                   35                  40                  45  
 Pro Gly Lys His Thr Phe Leu Cys Ser Lys His Phe Glu Ala Ser  
                   50                  55                  60  
 Cys Phe Asp Leu Thr Gly Gln Thr Arg Arg Leu Lys Met Asp Ala  
                   65                  70                  75  
 Val Pro Thr Ile Phe Asp Phe Cys Thr His Ile Lys Ser Met Lys  
                   80                  85                  90  
 Leu Lys Ser Arg Asn Leu Leu Lys Lys Asn Asn Ser Cys Ser Pro  
                   95                  100                  105  
 Ala Gly Pro Ser Asn Leu Lys Ser Asn Ile Ser Ser Gln Gln Val  
                   110                  115                  120  
 Leu Leu Glu His Ser Tyr Ala Phe Arg Asn Pro Met Glu Ala Lys  
                   125                  130                  135  
 Lys Arg Ile Ile Lys Leu Glu Lys Glu Ile Ala Ser Leu Arg Arg  
                   140                  145                  150  
 Lys Met Lys Thr Cys Leu Gln Lys Glu Arg Arg Ala Thr Arg Arg  
                   155                  160                  165  
 Trp Ile Lys Ala Thr Cys Leu Val Lys Asn Leu Glu Ala Asn Ser  
                   170                  175                  180  
 Val Leu Pro Lys Gly Thr Ser Glu His Met Leu Pro Thr Ala Leu  
                   185                  190                  195  
 Ser Ser Leu Pro Leu Glu Asp Phe Lys Ile Leu Glu Gln Asp Gln  
                   200                  205                  210  
 Gln Asp Lys Thr Leu Leu Ser Leu Asn Leu Lys Gln Thr Lys Ser  
                   215                  220                  225  
 Thr Phe Ile

<210> 18  
 <211> 503  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone Number: 1384286

<400> 18  
 Met Ala Thr Thr Val Thr Cys Thr Arg Phe Thr Asp Glu Tyr Gln  
           1                  5                  10                  15  
 Leu Tyr Glu Asp Ile Gly Lys Gly Ala Phe Ser Val Val Arg Arg  
                   20                  25                  30  
 Cys Val Lys Leu Cys Thr Gly His Glu Tyr Ala Ala Lys Ile Ile  
                   35                  40                  45  
 Asn Thr Lys Lys Leu Ser Ala Arg Asp His Gln Lys Leu Glu Arg

	50		55		60
Glu Ala Arg Ile	Cys Arg Leu Leu Lys	His Ser Asn Ile Val Arg			
	65		70		75
Leu His Asp Ser	Ile Ser Glu Glu Gly Phe	His Tyr Leu Val Phe			
	80		85		90
Asp Leu Val Thr	Gly Gly Glu Leu Phe Glu	Asp Ile Val Ala Arg			
	95		100		105
Glu Tyr Tyr Ser	Glu Ala Asp Ala Ser His	Cys Ile Gln Gln Ile			
	110		115		120
Leu Glu Ala Val	Leu His Cys His Gln Met	Gly Val Val His Arg			
	125		130		135
Asp Leu Lys Pro	Glu Asn Leu Leu Leu Ala	Ser Lys Cys Lys Gly			
	140		145		150
Ala Ala Val Lys	Leu Ala Asp Phe Gly Leu	Ala Ile Glu Val Gln			
	155		160		165
Gly Asp Gln Gln	Ala Trp Phe Gly Phe Ala	Gly Thr Pro Gly Tyr			
	170		175		180
Leu Ser Pro Glu	Val Leu Arg Lys Glu Ala	Tyr Gly Lys Pro Val			
	185		190		195
Asp Ile Trp Ala	Cys Gly Val Ile Leu Tyr	Ile Leu Leu Val Gly			
	200		205		210
Tyr Pro Pro Phe	Trp Asp Glu Asp Gln His	Lys Leu Tyr Gln Gln			
	215		220		225
Ile Lys Ala Gly	Ala Tyr Asp Phe Pro Ser	Pro Glu Trp Asp Thr			
	230		235		240
Val Thr Pro Glu	Ala Lys Asn Leu Ile Asn	Gln Met Leu Thr Ile			
	245		250		255
Asn Pro Ala Lys	Arg Ile Thr Ala His Glu	Ala Leu Lys His Pro			
	260		265		270
Trp Val Cys Gln	Arg Ser Thr Val Ala Ser	Met Met His Arg Gln			
	275		280		285
Glu Thr Val Glu	Cys Leu Lys Lys Phe Asn	Ala Arg Arg Lys Leu			
	290		295		300
Lys Gly Ala Ile	Leu Thr Thr Met Leu Ala	Thr Arg Asn Phe Ser			
	305		310		315
Ala Ala Lys Ser	Leu Leu Asn Lys Lys Ala	Asp Gly Val Lys Pro			
	320		325		330
His Thr Asn Ser	Thr Lys Asn Ser Ala Ala	Ala Thr Ser Pro Lys			
	335		340		345
Gly Thr Leu Pro	Pro Ala Ala Leu Glu Ser	Ser Asp Ser Ala Asn			
	350		355		360
Thr Thr Ile Glu	Asp Glu Asp Ala Lys Ala	Arg Lys Gln Glu Ile			
	365		370		375
Ile Lys Thr Thr	Glu Gln Leu Ile Glu Ala	Val Asn Asn Gly Asp			
	380		385		390
Phe Glu Ala Tyr	Ala Lys Ile Cys Asp Pro	Gly Leu Thr Ser Phe			
	395		400		405
Glu Pro Glu Ala	Leu Gly Asn Leu Val Glu	Gly Met Asp Phe His			
	410		415		420
Arg Phe Tyr Phe	Glu Asn Leu Leu Ala Lys	Asn Ser Lys Pro Ile			
	425		430		435
His Thr Thr Ile	Leu Asn Pro His Val His	Val Ile Gly Glu Asp			
	440		445		450
Ala Ala Cys Ile	Ala Tyr Ile Arg Leu Thr	Gln Tyr Ile Asp Gly			
	455		460		465
Gln Gly Arg Pro	Arg Thr Ser Gln Ser Glu	Glu Thr Arg Val Trp			

	470		475		480
His Arg Arg Asp	Gly Lys Trp Gln Asn Val	His Phe His Cys Ser			
	485		490		495
Gly Ala Pro Val	Ala Pro Leu Gln				
	500				

&lt;210&gt; 19

&lt;211&gt; 433

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 1512656

&lt;400&gt; 19

Met Thr Gly Glu Ala Gln Ala Gly Arg Lys Arg Ser Arg Ala Arg		
1	5	10 15
Pro Glu Gly Thr Glu Pro Val Arg Arg Glu Arg Thr Gln Pro Gly	20 25 30	
Leu Gly Pro Gly Arg Ala Arg Ala Met Ala Ala Glu Ala Thr Ala	35 40 45	
Val Ala Gly Ser Gly Ala Val Gly Gly Cys Leu Ala Lys Asp Gly	50 55 60	
Leu Gln Gln Ser Lys Cys Pro Asp Thr Thr Pro Lys Arg Arg Arg	65 70 75	
Ala Ser Ser Leu Ser Arg Asp Ala Glu Arg Arg Ala Tyr Gln Trp	80 85 90	
Cys Arg Glu Tyr Leu Gly Gly Ala Trp Arg Arg Val Gln Pro Glu	95 100 105	
Glu Leu Arg Val Tyr Pro Val Ser Gly Gly Leu Ser Asn Leu Leu	110 115 120	
Phe Arg Cys Ser Leu Pro Asp His Leu Pro Ser Val Gly Glu Glu	125 130 135	
Pro Arg Glu Val Leu Leu Arg Leu Tyr Gly Ala Ile Leu Gln Gly	140 145 150	
Val Asp Ser Leu Val Leu Glu Ser Val Met Phe Ala Ile Leu Ala	155 160 165	
Glu Arg Ser Leu Gly Pro Gln Leu Tyr Gly Val Phe Pro Glu Gly	170 175 180	
Arg Leu Glu Gln Tyr Ile Pro Ser Arg Pro Leu Lys Thr Gln Glu	185 190 195	
Leu Arg Glu Pro Val Leu Ser Ala Ala Ile Ala Thr Lys Met Ala	200 205 210	
Gln Phe His Gly Met Glu Met Pro Phe Thr Lys Glu Pro His Trp	215 220 225	
Leu Phe Gly Thr Met Glu Arg Tyr Leu Lys Gln Ile Gln Asp Leu	230 235 240	
Pro Pro Thr Gly Leu Pro Glu Met Asn Leu Leu Glu Met Tyr Ser	245 250 255	
Leu Lys Asp Glu Met Gly Asn Leu Arg Lys Leu Leu Glu Ser Thr	260 265 270	
Pro Ser Pro Val Val Phe Cys His Asn Asp Ile Gln Glu Gly Asn		

	275		280		285
Ile Leu Leu Leu	Ser Glu Pro Glu Asn	Ala Asp Ser Leu Met	Leu		
	290		295		300
Val Asp Phe Glu Tyr	Ser Ser Tyr Asn Tyr	Arg Gly Phe Asp	Ile		
	305		310		315
Gly Asn His Phe Cys	Glu Trp Val Tyr Asp	Tyr Thr His Glu	Glu		
	320		325		330
Trp Pro Phe Tyr Lys	Ala Arg Pro Thr Asp	Tyr Pro Thr Gln	Glu		
	335		340		345
Gln Gln Leu His Phe	Ile Arg His Tyr Leu	Ala Glu Ala Lys	Lys		
	350		355		360
Gly Glu Thr Leu Ser	Gln Glu Glu Gln Arg	Lys Leu Glu Glu	Asp		
	365		370		375
Leu Leu Val Glu Val	Ser Arg Tyr Ala Leu	Ala Ser His Phe	Phe		
	380		385		390
Trp Gly Leu Trp Ser	Ile Leu Gln Ala Ser	Met Ser Thr Ile	Glu		
	395		400		405
Phe Gly Tyr Leu Asp	Tyr Ala Gln Ser Arg	Phe Gln Phe Tyr	Phe		
	410		415		420
Gln Gln Lys Gly Gln	Leu Thr Ser Val His	Ser Ser Ser			
	425		430		

&lt;210&gt; 20

&lt;211&gt; 527

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2098635

&lt;400&gt; 20

Met Ser Leu Cys Gly	Ala Arg Ala Asn	Ala Lys Met Met	Ala Ala
1	5	10	15
Tyr Asn Gly Gly Thr	Ser Ala Ala Ala	Ala Gly His His	His His
	20	25	30
His His His His Leu	Pro His Leu Pro	Pro Pro His Leu	Leu His
	35	40	45
His His His Pro Gln	His His Leu His	Pro Gly Ser Ala	Ala Ala
	50	55	60
Val His Pro Val Gln	Gln His Thr Ser	Ser Ala Ala Ala	Ala Ala
	65	70	75
Ala Ala Ala Ala Ala	Ala Ala Ala Met	Leu Asn Pro Gly	Gln Gln
	80	85	90
Gln Pro Tyr Phe Pro	Ser Pro Ala Pro	Gly Gln Ala Pro	Gly Pro
	95	100	105
Ala Ala Ala Ala Pro	Ala Gln Val Gln	Ala Ala Ala Ala	Ala Thr
	110	115	120
Val Lys Ala His His	His Gln His Ser	His His Pro Gln	Gln Gln
	125	130	135
Leu Asp Ile Glu Pro	Asp Arg Pro Ile	Gly Tyr Gly Ala	Phe Gly
	140	145	150
Val Val Trp Ser Val	Thr Asp Pro Arg	Asp Gly Lys Arg	Val Ala
	155	160	165
Leu Lys Lys Met Pro	Asn Val Phe Gln	Asn Leu Val Ser	Cys Lys



	170		175		180
Arg Val Phe Arg	Glu Leu Lys Met Leu Cys Phe Phe Lys His Asp				
	185		190		195
Asn Val Leu Ser	Ala Leu Asp Ile Leu Gln Pro Pro His Ile Asp				
	200		205		210
Tyr Phe Glu Glu	Ile Tyr Val Val Thr Glu Leu Met Gln Ser Asp				
	215		220		225
Leu His Lys Ile	Ile Val Ser Pro Gln Pro Leu Ser Ser Asp His				
	230		235		240
Val Lys Val Phe	Leu Tyr Gln Ile Leu Arg Gly Leu Lys Tyr Leu				
	245		250		255
His Ser Ala Gly	Ile Leu His Arg Asp Ile Lys Pro Gly Asn Leu				
	260		265		270
Leu Val Asn Ser	Asn Cys Val Leu Lys Ile Cys Asp Phe Gly Leu				
	275		280		285
Ala Arg Val Glu	Glu Leu Asp Glu Ser Arg His Met Thr Gln Glu				
	290		295		300
Val Val Thr Gln	Tyr Tyr Arg Ala Pro Glu Ile Leu Met Gly Ser				
	305		310		315
Arg His Tyr Ser	Asn Ala Ile Asp Ile Trp Ser Val Gly Cys Ile				
	320		325		330
Phe Ala Glu Leu	Leu Gly Arg Arg Ile Leu Phe Gln Ala Gln Ser				
	335		340		345
Pro Ile Gln Gln	Leu Asp Leu Ile Thr Asp Leu Leu Gly Thr Pro				
	350		355		360
Ser Leu Glu Ala	Met Arg Thr Ala Cys Glu Gly Ala Lys Ala His				
	365		370		375
Ile Leu Arg Gly	Pro His Lys Gln Pro Ser Leu Pro Val Leu Tyr				
	380		385		390
Thr Leu Ser Ser	Gln Ala Thr His Glu Ala Val His Leu Leu Cys				
	395		400		405
Arg Met Leu Val	Phe Asp Pro Ser Lys Arg Ile Ser Ala Lys Asp				
	410		415		420
Ala Leu Ala His	Pro Tyr Leu Asp Glu Gly Arg Leu Arg Tyr His				
	425		430		435
Thr Cys Met Cys	Lys Cys Cys Phe Ser Thr Ser Thr Gly Arg Val				
	440		445		450
Tyr Thr Ser Asp	Phe Glu Pro Val Thr Asn Pro Lys Phe Asp Asp				
	455		460		465
Thr Phe Glu Lys	Asn Leu Ser Ser Val Arg Gln Val Lys Glu Ile				
	470		475		480
Ile His Gln Phe	Ile Leu Glu Gln Gln Lys Gly Asn Arg Val Pro				
	485		490		495
Leu Cys Ile Asn	Pro Gln Ser Ala Ala Phe Lys Ser Phe Ile Ser				
	500		505		510
Ser Thr Val Ala	Gln Pro Ser Glu Met Pro Pro Ser Pro Leu Val				
	515		520		525
Trp Glu					

&lt;210&gt; 21

&lt;211&gt; 322

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2446646

&lt;400&gt; 21

```

Met Glu Gly Ile Ser Asn Phe Lys Thr Pro Ser Lys Leu Ser Glu
 1           5           10           15
Lys Lys Lys Ser Val Leu Cys Ser Thr Pro Thr Ile Asn Ile Pro
           20           25           30
Ala Ser Pro Phe Met Gln Lys Leu Gly Phe Gly Thr Gly Val Asn
           35           40           45
Val Tyr Leu Met Lys Arg Ser Pro Arg Gly Leu Ser His Ser Pro
           50           55           60
Trp Ala Val Lys Lys Ile Asn Pro Ile Cys Asn Asp His Tyr Arg
           65           70           75
Ser Val Tyr Gln Lys Arg Leu Met Asp Glu Ala Lys Ile Leu Lys
           80           85           90
Ser Leu His His Pro Asn Ile Val Gly Tyr Arg Ala Phe Thr Glu
           95          100          105
Ala Asn Asp Gly Ser Leu Cys Leu Ala Met Glu Tyr Gly Gly Glu
          110          115          120
Lys Ser Leu Asn Asp Leu Ile Glu Glu Arg Tyr Lys Ala Ser Gln
          125          130          135
Asp Pro Phe Pro Ala Ala Ile Ile Leu Lys Val Ala Leu Asn Met
          140          145          150
Ala Arg Gly Leu Lys Tyr Leu His Gln Glu Lys Lys Leu Leu His
          155          160          165
Gly Asp Ile Lys Ser Ser Asn Val Val Ile Lys Gly Asp Phe Glu
          170          175          180
Thr Ile Lys Ile Cys Asp Val Gly Val Ser Leu Pro Leu Asp Glu
          185          190          195
Asn Met Thr Val Thr Asp Pro Glu Ala Cys Tyr Ile Gly Thr Glu
          200          205          210
Pro Trp Lys Pro Lys Glu Ala Val Glu Glu Asn Gly Val Ile Thr
          215          220          225
Asp Lys Ala Asp Ile Phe Ala Phe Gly Leu Thr Leu Trp Glu Met
          230          235          240
Met Thr Leu Ser Ile Pro His Ile Asn Leu Ser Asn Asp Asp Asp
          245          250          255
Asp Glu Asp Lys Thr Phe Asp Glu Ser Asp Phe Asp Asp Glu Ala
          260          265          270
Tyr Tyr Ala Ala Leu Gly Thr Arg Pro Pro Ile Asn Met Glu Glu
          275          280          285
Leu Asp Glu Ser Tyr Gln Lys Val Ile Glu Leu Phe Ser Val Cys
          290          295          300
Thr Asn Glu Asp Pro Lys Asp Arg Pro Ser Ala Ala His Ile Val
          305          310          315
Glu Ala Leu Glu Thr Asp Val
          320

```

&lt;210&gt; 22

&lt;211&gt; 802

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2764911

&lt;400&gt; 22

Met	Glu	Glu	Glu	Gly	Gly	Ser	Ser	Gly	Gly	Ala	Ala	Gly	Thr	Ser	1	5	10	15
Ala	Asp	Gly	Gly	Asp	Gly	Gly	Glu	Gln	Leu	Leu	Thr	Val	Lys	His	20	25	30	
Glu	Leu	Arg	Thr	Ala	Asn	Leu	Thr	Gly	His	Ala	Glu	Lys	Val	Gly	35	40	45	
Ile	Glu	Asn	Phe	Glu	Leu	Leu	Lys	Val	Leu	Gly	Thr	Gly	Ala	Tyr	50	55	60	
Gly	Lys	Val	Phe	Leu	Val	Arg	Lys	Ile	Ser	Gly	His	Asp	Thr	Gly	65	70	75	
Lys	Leu	Tyr	Ala	Met	Lys	Val	Leu	Lys	Lys	Ala	Thr	Ile	Val	Gln	80	85	90	
Lys	Ala	Lys	Thr	Thr	Glu	His	Thr	Arg	Thr	Glu	Arg	Gln	Val	Leu	95	100	105	
Glu	His	Ile	Arg	Gln	Ser	Pro	Phe	Leu	Val	Thr	Leu	His	Tyr	Ala	110	115	120	
Phe	Gln	Thr	Glu	Thr	Lys	Leu	His	Leu	Ile	Leu	Asp	Tyr	Ile	Asn	125	130	135	
Gly	Gly	Glu	Leu	Phe	Thr	His	Leu	Ser	Gln	Arg	Glu	Arg	Phe	Thr	140	145	150	
Glu	His	Glu	Val	Gln	Ile	Tyr	Val	Gly	Glu	Ile	Val	Leu	Ala	Leu	155	160	165	
Glu	His	Leu	His	Lys	Leu	Gly	Ile	Ile	Tyr	Arg	Asp	Ile	Lys	Leu	170	175	180	
Glu	Asn	Ile	Leu	Leu	Asp	Ser	Asn	Gly	His	Val	Val	Leu	Thr	Asp	185	190	195	
Phe	Gly	Leu	Ser	Lys	Glu	Phe	Val	Ala	Asp	Glu	Thr	Glu	Arg	Ala	200	205	210	
Tyr	Ser	Phe	Cys	Gly	Thr	Ile	Glu	Tyr	Met	Ala	Pro	Asp	Ile	Val	215	220	225	
Arg	Gly	Gly	Asp	Ser	Gly	His	Asp	Lys	Ala	Val	Asp	Trp	Trp	Ser	230	235	240	
Leu	Gly	Val	Leu	Met	Tyr	Glu	Leu	Leu	Thr	Gly	Ala	Ser	Pro	Phe	245	250	255	
Thr	Val	Asp	Gly	Glu	Lys	Asn	Ser	Gln	Ala	Glu	Ile	Ser	Arg	Arg	260	265	270	
Ile	Leu	Lys	Ser	Glu	Pro	Pro	Tyr	Pro	Gln	Glu	Met	Ser	Ala	Leu	275	280	285	
Ala	Lys	Asp	Leu	Ile	Gln	Arg	Leu	Leu	Met	Lys	Asp	Pro	Lys	Lys	290	295	300	
Arg	Leu	Gly	Cys	Gly	Pro	Arg	Asp	Ala	Asp	Glu	Ile	Lys	Glu	His	305	310	315	
Leu	Phe	Phe	Gln	Lys	Ile	Asn	Trp	Asp	Asp	Leu	Ala	Ala	Lys	Lys	320	325	330	
Val	Pro	Ala	Pro	Phe	Lys	Pro	Val	Ile	Arg	Asp	Glu	Leu	Asp	Val	335	340	345	
Ser	Asn	Phe	Ala	Glu	Phe	Thr	Glu	Met	Asp	Pro	Thr	Tyr	Ser		350	355	360	
Pro	Ala	Ala	Leu	Pro	Gln	Ser	Ser	Glu	Lys	Leu	Phe	Gln	Gly	Tyr	365	370	375	
Ser	Phe	Val	Ala	Pro	Ser	Ile	Leu	Phe	Lys	Arg	Asn	Ala	Ala	Val	380	385	390	
Ile	Asp	Pro	Leu	Gln	Phe	His	Met	Gly	Val	Glu	Arg	Pro	Gly	Val				

	395		400		405
Thr Asn Val Ala	Arg Ser Ala Met Met	Lys Asp Ser Pro Phe Tyr			
	410		415		420
Gln His Tyr Asp	Leu Asp Leu Lys Asp	Lys Pro Leu Gly Glu Gly			
	425		430		435
Ser Phe Ser Ile	Cys Arg Lys Cys Val	His Lys Lys Ser Asn Gln			
	440		445		450
Ala Phe Ala Val	Lys Ile Ile Ser Lys	Arg Met Glu Ala Asn Thr			
	455		460		465
Gln Lys Glu Ile	Thr Ala Leu Glu Leu	Cys Glu Gly His Pro Asn			
	470		475		480
Ile Val Lys Leu	His Glu Val Phe His	Asp Gln Leu His Thr Phe			
	485		490		495
Leu Val Met Glu	Leu Leu Asn Gly Gly	Glu Leu Phe Glu Arg Ile			
	500		505		510
Lys Lys Lys Lys	His Phe Ser Glu Thr	Glu Ala Ser Tyr Ile Met			
	515		520		525
Arg Lys Leu Val	Ser Ala Val Ser His	Met His Asp Val Gly Val			
	530		535		540
Val His Arg Asp	Leu Lys Pro Glu Asn	Leu Leu Phe Thr Asp Glu			
	545		550		555
Asn Asp Asn Leu	Glu Ile Lys Ile Ile	Asp Phe Gly Phe Ala Arg			
	560		565		570
Leu Lys Pro Pro	Asp Asn Gln Pro Leu	Lys Thr Pro Cys Phe Thr			
	575		580		585
Leu His Tyr Ala	Ala Pro Glu Leu Leu	Asn Gln Asn Gly Tyr Asp			
	590		595		600
Glu Ser Cys Asp	Leu Trp Ser Leu Gly	Val Ile Leu Tyr Thr Met			
	605		610		615
Leu Ser Gly Gln	Val Pro Phe Gln Ser	His Asp Arg Ser Leu Thr			
	620		625		630
Cys Thr Ser Ala	Val Glu Ile Met Lys	Lys Ile Lys Lys Gly Asp			
	635		640		645
Phe Ser Phe Glu	Gly Glu Ala Trp Lys	Asn Val Ser Gln Glu Ala			
	650		655		660
Lys Asp Leu Ile	Gln Gly Leu Leu Thr	Val Asp Pro Asn Lys Arg			
	665		670		675
Leu Lys Met Ser	Gly Leu Arg Tyr Asn	Glu Trp Leu Gln Asp Gly			
	680		685		690
Ser Gln Leu Ser	Ser Asn Pro Leu Met	Thr Pro Asp Ile Leu Gly			
	695		700		705
Ser Ser Gly Ala	Ala Val His Thr Cys	Val Lys Ala Thr Phe His			
	710		715		720
Ala Phe Asn Lys	Tyr Lys Arg Glu Gly	Phe Cys Leu Gln Asn Val			
	725		730		735
Asp Lys Ala Pro	Leu Ala Lys Arg Arg	Lys Met Lys Lys Thr Ser			
	740		745		750
Thr Ser Thr Glu	Thr Arg Ser Ser Ser	Ser Glu Ser Ser His Ser			
	755		760		765
Ser Ser Ser His	Ser His Gly Lys Thr	Thr Pro Thr Lys Thr Leu			
	770		775		780
Gln Pro Ser Asn	Pro Ala Asp Ser Asn	Asn Pro Glu Thr Leu Phe			
	785		790		795
Gln Phe Ser Asp	Ser Val Ala				
	800				

<210> 23  
 <211> 641  
 <212> PRT  
 <213> Homo sapiens  
  
 <220>  
 <221> misc\_feature  
 <223> Incyte Clone Number: 3013946

<400> 23  
 Met Ala Thr Thr Val Thr Cys Thr Arg Phe Thr Asp Glu Tyr Gln  
 1 5 10 15  
 Leu Tyr Glu Asp Ile Gly Lys Gly Ala Phe Ser Val Val Arg Arg  
 20 25 30  
 Cys Val Lys Leu Cys Thr Gly His Glu Tyr Ala Ala Lys Ile Ile  
 35 40 45  
 Asn Thr Lys Lys Leu Ser Ala Arg Asp His Gln Lys Leu Glu Arg  
 50 55 60  
 Glu Ala Arg Ile Cys Arg Leu Leu Lys His Ser Asn Ile Val Arg  
 65 70 75  
 Leu His Asp Ser Ile Ser Glu Glu Gly Phe His Tyr Leu Val Phe  
 80 85 90  
 Asp Leu Val Thr Gly Gly Glu Leu Phe Glu Asp Ile Val Ala Arg  
 95 100 105  
 Glu Tyr Tyr Ser Glu Ala Asp Ala Ser His Cys Ile Gln Gln Ile  
 110 115 120  
 Leu Glu Ala Val Leu His Cys His Gln Met Gly Val Val His Arg  
 125 130 135  
 Asp Leu Lys Pro Glu Asn Leu Leu Leu Ala Ser Lys Cys Lys Gly  
 140 145 150  
 Ala Ala Val Lys Leu Ala Asp Phe Gly Leu Ala Ile Glu Val Gln  
 155 160 165  
 Gly Asp Gln Gln Ala Trp Phe Gly Phe Ala Gly Thr Pro Gly Tyr  
 170 175 180  
 Leu Ser Pro Glu Val Leu Arg Lys Glu Ala Tyr Gly Lys Pro Val  
 185 190 195  
 Asp Ile Trp Ala Cys Gly Val Ile Leu Tyr Ile Leu Leu Val Gly  
 200 205 210  
 Tyr Pro Pro Phe Trp Asp Glu Asp Gln His Lys Leu Tyr Gln Gln  
 215 220 225  
 Ile Lys Ala Gly Ala Tyr Asp Phe Pro Ser Pro Glu Trp Asp Thr  
 230 235 240  
 Val Thr Pro Glu Ala Lys Asn Leu Ile Asn Gln Met Leu Thr Ile  
 245 250 255  
 Asn Pro Ala Lys Arg Ile Thr Ala His Glu Ala Leu Lys His Pro  
 260 265 270  
 Trp Val Cys Gln Arg Ser Thr Val Ala Ser Met Met His Arg Gln  
 275 280 285  
 Glu Thr Val Glu Cys Leu Lys Lys Phe Asn Ala Arg Arg Lys Leu  
 290 295 300  
 Lys Gly Ala Ile Leu Thr Thr Met Leu Ala Thr Arg Asn Phe Ser  
 305 310 315  
 Ala Lys Ser Leu Leu Asn Lys Lys Ala Asp Gly Val Lys Pro Gln  
 320 325 330  
 Thr Asn Ser Thr Lys Asn Ser Ala Ala Ala Thr Ser Pro Lys Gly  
 335 340 345

```

Thr Leu Pro Pro Ala Ala Leu Glu Pro Gln Thr Thr Val Ile His
350 355 360
Asn Pro Val Asp Gly Ile Lys Glu Ser Ser Asp Ser Ala Asn Thr
365 370 375
Thr Ile Glu Asp Glu Asp Ala Lys Ala Pro Arg Val Pro Asp Ile
380 385 390
Leu Ser Ser Val Arg Arg Gly Ser Gly Ala Pro Glu Ala Glu Gly
395 400 405
Pro Leu Pro Cys Pro Ser Pro Ala Pro Phe Gly Pro Leu Pro Ala
410 415 420
Pro Ser Pro Arg Ile Ser Asp Ile Leu Asn Ser Val Arg Arg Gly
425 430 435
Ser Gly Thr Pro Glu Ala Glu Gly Pro Leu Ser Ala Gly Pro Pro
440 445 450
Pro Cys Leu Ser Pro Ala Leu Leu Gly Pro Leu Ser Ser Pro Ser
455 460 465
Pro Arg Ile Ser Asp Ile Leu Asn Ser Val Arg Arg Gly Ser Gly
470 475 480
Thr Pro Glu Ala Lys Gly Pro Ser Pro Val Gly Pro Pro Pro Cys
485 490 495
Pro Ser Pro Thr Ile Pro Gly Pro Leu Pro Thr Pro Ser Arg Lys
500 505 510
Gln Glu Ile Ile Lys Thr Thr Glu Gln Leu Ile Glu Ala Val Asn
515 520 525
Asn Gly Asp Phe Glu Ala Tyr Ala Lys Ile Cys Asp Pro Gly Leu
530 535 540
Thr Ser Phe Glu Pro Glu Ala Leu Gly Asn Leu Val Glu Gly Met
545 550 555
Asp Phe His Arg Phe Tyr Phe Glu Asn Leu Leu Ala Lys Asn Ser
560 565 570
Lys Pro Ile His Thr Thr Ile Leu Asn Pro His Val His Val Ile
575 580 585
Gly Glu Asp Ala Ala Cys Ile Ala Tyr Ile Arg Leu Thr Gln Tyr
590 595 600
Ile Asp Gly Gln Gly Arg Pro Arg Thr Ser Gln Ser Glu Glu Thr
605 610 615
Arg Val Trp His Arg Arg Asp Gly Lys Trp Gln Asn Val His Phe
620 625 630
His Cys Ser Gly Ala Pro Val Ala Pro Leu Gln
635 640

```

&lt;210&gt; 24

&lt;211&gt; 588

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 067967

&lt;400&gt; 24

```

Met Gly Gly Thr Ala Arg Gly Pro Gly Arg Lys Asp Ala Gly Pro
1 5 10 15
Pro Gly Ala Gly Leu Pro Pro Gln Gln Arg Arg Leu Gly Asp Gly
20 25 30
Val Tyr Asp Thr Phe Met Met Ile Asp Glu Thr Lys Cys Pro Pro

```

	35		40		45
Cys Ser Asn Val	Leu Cys Asn Pro Ser	Glu Pro Pro Ser Pro Arg			
	50		55		60
Arg Leu Asn Met	Thr Thr Glu Gln Phe	Thr Gly Asp His Thr Gln			
	65		70		75
His Phe Leu Asp	Gly Gly Glu Met Lys	Val Glu Gln Leu Phe Gln			
	80		85		90
Glu Phe Gly Asn	Arg Lys Ser Asn Thr	Ile Gln Ser Asp Gly Ile			
	95		100		105
Ser Asp Ser Glu	Lys Cys Ser Pro Thr	Val Ser Gln Gly Lys Ser			
	110		115		120
Ser Asp Cys Leu	Asn Thr Val Lys Ser	Asn Ser Ser Ser Lys Ala			
	125		130		135
Pro Lys Val Val	Pro Leu Thr Pro Glu	Gln Ala Leu Lys Gln Tyr			
	140		145		150
Lys His His Leu	Thr Ala Tyr Glu Lys	Leu Glu Ile Ile Asn Tyr			
	155		160		165
Pro Glu Ile Tyr	Phe Val Gly Pro Asn	Ala Lys Lys Arg His Gly			
	170		175		180
Val Ile Gly Gly	Pro Asn Asn Gly Gly	Tyr Asp Asp Ala Asp Gly			
	185		190		195
Ala Tyr Ile His	Val Pro Arg Asp His	Leu Ala Tyr Arg Tyr Glu			
	200		205		210
Val Leu Lys Ile	Ile Gly Lys Gly Ser	Phe Gly Gln Val Ala Arg			
	215		220		225
Val Tyr Asp His	Lys Leu Arg Gln Tyr	Val Ala Leu Lys Met Val			
	230		235		240
Arg Asn Glu Lys	Arg Phe His Arg Gln	Ala Ala Glu Glu Ile Arg			
	245		250		255
Ile Leu Glu His	Leu Lys Lys Gln Asp	Lys Thr Gly Ser Met Asn			
	260		265		270
Val Ile His Met	Leu Glu Ser Phe Thr	Phe Arg Asn His Val Cys			
	275		280		285
Met Ala Phe Glu	Leu Leu Ser Ile Asp	Leu Tyr Glu Leu Ile Lys			
	290		295		300
Lys Asn Lys Phe	Gln Gly Phe Ser Val	Gln Leu Val Arg Lys Phe			
	305		310		315
Ala Gln Ser Ile	Leu Gln Ser Leu Asp	Ala Leu His Lys Asn Lys			
	320		325		330
Ile Ile His Cys	Asp Leu Lys Pro Glu	Asn Ile Leu Leu Lys His			
	335		340		345
His Gly Arg Ser	Ser Thr Lys Val Ile	Asp Phe Gly Ser Ser Cys			
	350		355		360
Phe Glu Tyr Gln	Lys Leu Tyr Thr Tyr	Ile Gln Ser Arg Phe Tyr			
	365		370		375
Arg Ala Pro Glu	Ile Ile Leu Gly Ser	Arg Tyr Ser Thr Pro Ile			
	380		385		390
Asp Ile Trp Ser	Phe Gly Cys Ile Leu	Ala Glu Leu Leu Thr Gly			
	395		400		405
Gln Pro Leu Phe	Pro Gly Glu Asp Glu	Gly Asp Gln Leu Ala Cys			
	410		415		420
Met Met Glu Leu	Leu Gly Met Pro Pro	Pro Lys Leu Leu Glu Gln			
	425		430		435
Ser Lys Arg Ala	Lys Tyr Phe Ile Asn	Ser Lys Gly Ile Pro Arg			
	440		445		450
Tyr Cys Ser Val	Thr Thr Gln Ala Asp	Gly Arg Val Val Leu Val			

				455					460					465
Gly	Gly	Arg	Ser	Arg	Arg	Gly	Lys	Lys	Arg	Gly	Pro	Pro	Gly	Ser
				470					475					480
Lys	Asp	Trp	Gly	Thr	Ala	Leu	Lys	Gly	Cys	Asp	Asp	Tyr	Leu	Phe
				485					490					495
Ile	Glu	Phe	Leu	Lys	Arg	Cys	Leu	His	Trp	Asp	Pro	Ser	Ala	Arg
				500					505					510
Leu	Thr	Pro	Ala	Gln	Ala	Leu	Arg	His	Pro	Trp	Ile	Ser	Lys	Ser
				515					520					525
Val	Pro	Arg	Pro	Leu	Thr	Thr	Ile	Asp	Lys	Val	Ser	Gly	Lys	Arg
				530					535					540
Val	Val	Asn	Pro	Ala	Ser	Ala	Phe	Gln	Gly	Leu	Gly	Ser	Lys	Leu
				545					550					555
Pro	Pro	Val	Val	Gly	Ile	Ala	Asn	Lys	Leu	Lys	Ala	Asn	Leu	Met
				560					565					570
Ser	Glu	Thr	Asn	Gly	Ser	Ile	Pro	Leu	Cys	Ser	Val	Leu	Pro	Lys
				575					580					585
Leu	Ile	Ser												

&lt;210&gt; 25

&lt;211&gt; 389

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 346275

&lt;400&gt; 25

Met	Ser	Asp	Val	Cys	Ser	Ser	Gln	Arg	Ala	Glu	His	Glu	His	Leu
1				5					10					15
Pro	Gly	Leu	Val	Pro	Pro	Pro	Ser	Gly	Met	Gly	Val	Arg	Lys	Gly
				20					25					30
Ser	Ser	Pro	Leu	Lys	Ser	His	Pro	Cys	Arg	Glu	Lys	Ser	Val	Ser
				35					40					45
Asn	Arg	Arg	Ser	Gly	Lys	Thr	Ile	Val	Arg	Ser	Ala	Val	Glu	Glu
				50					55					60
Val	Arg	Thr	Ala	Gly	Leu	Phe	Arg	Ser	Gly	Phe	Ser	Glu	Glu	Lys
				65					70					75
Ala	Thr	Gly	Lys	Leu	Phe	Ala	Val	Lys	Cys	Ile	Pro	Lys	Lys	Ala
				80					85					90
Leu	Lys	Gly	Lys	Glu	Ser	Ser	Ile	Glu	Asn	Glu	Ile	Ala	Val	Leu
				95					100					105
Arg	Lys	Ile	Lys	His	Glu	Asn	Ile	Val	Ala	Leu	Glu	Asp	Ile	Tyr
				110					115					120
Glu	Ser	Pro	Asn	His	Leu	Tyr	Leu	Val	Met	Gln	Leu	Val	Ser	Gly
				125					130					135
Gly	Glu	Leu	Phe	Asp	Arg	Ile	Val	Glu	Lys	Gly	Phe	Tyr	Thr	Glu
				140					145					150
Lys	Asp	Ala	Ser	Thr	Leu	Ile	Arg	Gln	Val	Leu	Asp	Ala	Val	Tyr
				155					160					165
Tyr	Leu	His	Arg	Met	Gly	Ile	Val	His	Arg	Asp	Leu	Lys	Pro	Glu
				170					175					180
Asn	Leu	Leu	Tyr	Tyr	Ser	Gln	Asp	Glu	Glu	Ser	Lys	Ile	Met	Ile
				185					190					195



Ser Asp Phe Gly Leu Ser Lys Met Glu Gly Lys Gly Asp Val Met		
	200	205 210
Ser Thr Ala Cys Gly Thr Pro Gly Tyr Val Ala Pro Glu Val Leu		
	215	220 225
Ala Gln Lys Pro Tyr Ser Lys Ala Val Asp Cys Trp Ser Ile Gly		
	230	235 240
Val Ile Ala Tyr Ile Leu Leu Cys Gly Tyr Pro Pro Phe Tyr Asp		
	245	250 255
Glu Asn Asp Ser Lys Leu Phe Glu Gln Ile Leu Lys Ala Glu Tyr		
	260	265 270
Glu Phe Asp Ser Pro Tyr Trp Asp Asp Ile Ser Asp Ser Ala Lys		
	275	280 285
Asp Phe Ile Arg Asn Leu Met Glu Lys Asp Pro Asn Lys Arg Tyr		
	290	295 300
Thr Cys Glu Gln Ala Ala Arg His Pro Trp Ile Ala Gly Asp Thr		
	305	310 315
Ala Leu Asn Lys Asn Ile His Glu Ser Val Ser Ala Gln Ile Arg		
	320	325 330
Lys Asn Phe Ala Lys Ser Lys Trp Arg Gln Ala Phe Asn Ala Thr		
	335	340 345
Ala Val Val Arg His Met Arg Lys Leu His Leu Gly Ser Ser Leu		
	350	355 360
Asp Ser Ser Asn Ala Ser Val Ser Ser Ser Leu Ser Leu Ala Ser		
	365	370 375
Gln Lys Asp Cys Ala Tyr Val Ala Lys Pro Glu Ser Leu Ser		
	380	385

&lt;210&gt; 26

&lt;211&gt; 343

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 283746

&lt;400&gt; 26

Met Ile Gly Glu Glu Ala Met Ile Asn Tyr Glu Asn Phe Leu Lys		
1	5	10 15
Val Gly Glu Lys Ala Gly Ala Lys Cys Lys Gln Phe Phe Thr Ala		
	20	25 30
Lys Val Phe Ala Lys Leu Leu His Thr Asp Ser Tyr Gly Arg Ile		
	35	40 45
Ser Ile Met Gln Phe Phe Asn Tyr Val Met Arg Lys Val Trp Leu		
	50	55 60
His Gln Thr Arg Ile Gly Leu Ser Leu Tyr Asp Val Ala Gly Gln		
	65	70 75
Gly Tyr Leu Arg Glu Ser Asp Leu Glu Asn Tyr Ile Leu Glu Leu		
	80	85 90
Ile Pro Thr Leu Pro Gln Leu Asp Gly Leu Glu Lys Ser Phe Tyr		
	95	100 105
Ser Phe Tyr Val Cys Thr Ala Val Arg Lys Phe Phe Phe Phe Leu		
	110	115 120
Asp Pro Leu Arg Thr Gly Lys Ile Lys Ile Gln Asp Ile Leu Ala		
	125	130 135

<400> 27															
Met	Gly	Asn	Gly	Met	Asn	Lys	Ile	Leu	Pro	Gly	Leu	Tyr	Ile	Gly	
1				5					10					15	
Asn	Phe	Lys	Asp	Ala	Arg	Asp	Ala	Glu	Gln	Leu	Ser	Lys	Asn	Lys	
				20					25					30	
Val	Thr	His	Ile	Leu	Ser	Val	His	Asp	Ser	Ala	Arg	Pro	Met	Leu	
				35					40					45	
Glu	Gly	Val	Lys	Tyr	Leu	Cys	Ile	Pro	Ala	Ala	Asp	Ser	Pro	Ser	
				50					55					60	
Gln	Asn	Leu	Thr	Arg	His	Phe	Lys	Glu	Ser	Ile	Lys	Phe	Ile	His	
				65					70					75	
Glu	Cys	Arg	Leu	Arg	Gly	Glu	Ser	Cys	Leu	Val	His	Cys	Leu	Ala	
				80					85					90	
Gly	Val	Ser	Arg	Ser	Val	Thr	Leu	Val	Ile	Ala	Tyr	Ile	Met	Thr	
				95					100					105	
Val	Thr	Asp	Phe	Gly	Trp	Glu	Asp	Ala	Leu	His	Thr	Val	Arg	Ala	
				110					115					120	

Gly	Arg	Ser	Cys	Ala	Asn	Pro	Asn	Val	Gly	Phe	Gln	Arg	Gln	Leu
				125					130					135
Gln	Glu	Phe	Glu	Lys	His	Glu	Val	His	Gln	Tyr	Arg	Gln	Trp	Leu
				140					145					150
Lys	Glu	Glu	Tyr	Gly	Glu	Ser	Pro	Leu	Gln	Asp	Ala	Glu	Glu	Ala
				155					160					165
Lys	Asn	Ile	Leu	Ala	Ala	Pro	Gly	Ile	Leu	Lys	Phe	Trp	Ala	Phe
				170					175					180
Leu	Arg	Arg	Leu											

<210> 29  
 <211> 118  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone Number: 619292

Met	Gly	Leu	Ile	Asp	Gly	Met	His	Thr	His	Leu	Gly	Ala	Pro	Gly
1				5					10					15
Leu	Tyr	Ile	Gln	Thr	Leu	Leu	Pro	Gly	Ser	Pro	Ala	Ala	Ala	Asp
				20					25					30
Gly	Arg	Leu	Ser	Leu	Gly	Asp	Arg	Ile	Leu	Glu	Val	Asn	Gly	Ser
				35					40					45
Ser	Leu	Leu	Gly	Leu	Gly	Tyr	Leu	Arg	Ala	Val	Asp	Leu	Ile	Arg
				50					55					60
His	Gly	Gly	Lys	Lys	Met	Arg	Phe	Leu	Val	Ala	Lys	Ser	Asp	Val
				65					70					75
Gly	Lys	Gln	Pro	Arg	Arg	Ser	Ile	Ser	Ala	Arg	Pro	Leu	Ser	Arg
				80					85					90
Gly	Ala	Ala	Arg	Thr	Pro	Pro	Gln	Ala	Arg	His	Pro	Val	Pro	Pro
				95					100					105
Gly	Asp	Thr	Gly	Leu	Pro	Pro	Ala	Phe	Val	Pro	Val	Leu		
				110					115					

<210> 30  
 <211> 356  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone Number: 2054049

Met	Val	Gly	Val	Ser	Gly	Lys	Arg	Ser	Lys	Glu	Asp	Glu	Lys	Tyr
1				5					10					15
Leu	Gln	Ala	Ile	Met	Asp	Ser	Asn	Ala	Gln	Ser	His	Lys	Ile	Phe
				20					25					30
Ile	Phe	Asp	Ala	Arg	Pro	Ser	Val	Asn	Ala	Val	Ala	Asn	Lys	Ala
				35					40					45
Lys	Gly	Gly	Gly	Tyr	Glu	Ser	Glu	Asp	Ala	Tyr	Gln	Asn	Ala	Glu

50	55	60
Leu Val Phe Leu Asp Ile His Asn Ile	His Val Met Arg Glu Ser	
65	70	75
Leu Arg Lys Leu Lys Glu Ile Val Tyr	Pro Asn Ile Glu Glu Thr	
80	85	90
His Trp Leu Ser Asn Leu Glu Ser Thr	His Trp Leu Glu His Ile	
95	100	105
Lys Leu Ile Leu Ala Gly Ala Leu Arg	Ile Ala Asp Lys Val Glu	
110	115	120
Ser Gly Lys Thr Ser Val Val Val His	Cys Ser Asp Gly Trp Asp	
125	130	135
Arg Thr Ala Gln Leu Thr Ser Leu Ala	Met Leu Met Leu Asp Gly	
140	145	150
Tyr Tyr Arg Thr Ile Arg Gly Phe Glu	Val Leu Val Glu Lys Glu	
155	160	165
Trp Leu Ser Phe Gly His Arg Phe Gln	Leu Arg Val Gly His Gly	
170	175	180
Asp Lys Asn His Ala Asp Ala Asp Arg	Ser Pro Val Phe Leu Gln	
185	190	195
Phe Ile Asp Cys Val Trp Gln Met Thr	Arg Gln Phe Pro Thr Ala	
200	205	210
Phe Glu Phe Asn Glu Tyr Phe Leu Ile	Thr Ile Leu Asp His Leu	
215	220	225
Tyr Ser Cys Leu Phe Gly Thr Phe Leu	Cys Asn Ser Glu Gln Gln	
230	235	240
Arg Gly Lys Glu Asn Leu Pro Lys Arg	Thr Val Ser Leu Trp Ser	
245	250	255
Tyr Ile Asn Ser Gln Leu Glu Asp Phe	Thr Asn Pro Leu Tyr Gly	
260	265	270
Ser Tyr Ser Asn His Val Leu Tyr Pro	Val Ala Ser Met Arg His	
275	280	285
Leu Glu Leu Trp Val Gly Tyr Tyr Ile	Arg Trp Asn Pro Arg Met	
290	295	300
Lys Pro Gln Glu Pro Ile His Asn Arg	Tyr Lys Glu Leu Leu Ala	
305	310	315
Lys Arg Ala Glu Leu Gln Lys Lys Val	Glu Glu Leu Gln Arg Glu	
320	325	330
Ile Ser Asn Arg Ser Thr Ser Ser Ser	Glu Arg Ala Ser Ser Pro	
335	340	345
Ala Gln Cys Val Thr Pro Val Gln Thr	Val Val	
350	355	

&lt;210&gt; 31

&lt;211&gt; 453

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2843910

&lt;400&gt; 31

Met Ala Gly Ala Gly Gly Phe Gly Cys	Pro Ala Gly Gly Asn Asp
1	5 10 15
Phe Gln Trp Cys Phe Ser Gln Val Lys	Gly Ala Ile Asp Glu Asp

	20		25		30
Val	Ala	Glu	Ala	Asp	Ile
	35		40		45
Gly	Asp	Leu	Leu	Ala	Thr
	50		55		60
Phe	Gln	Arg	Glu	Gln	Glu
	65		70		75
Glu	Tyr	Asn	Val	Tyr	Ser
	80		85		90
Asp	Tyr	Leu	Lys	Ser	Leu
	95		100		105
Arg	Trp	Leu	Pro	Gln	Gln
	110		115		120
Asn	Asp	Lys	Thr	Ile	Lys
	125		130		135
Arg	Ala	Glu	Gly	Tyr	Asn
	140		145		150
Asp	Pro	Phe	Arg	Ile	Thr
	155		160		165
Met	Asp	Leu	Met	Val	Glu
	170		175		180
Ala	His	Thr	Tyr	His	Ile
	185		190		195
Glu	Thr	Tyr	Leu	Ser	Ala
	200		205		210
Leu	Glu	Ile	Thr	Asp	Arg
	215		220		225
Ala	Asn	Met	Glu	Glu	Leu
	230		235		240
His	Pro	His	Gln	Cys	Asn
	245		250		255
Thr	Ile	Arg	Leu	Cys	Asp
	260		265		270
His	Ser	Lys	Phe	Phe	Glu
	275		280		285
Phe	Phe	Ser	Glu	Ile	Ile
	290		295		300
His	Ser	Gly	Arg	Tyr	Met
	305		310		315
Val	Trp	Asp	Leu	Asn	Met
	320		325		330
Val	His	Glu	Tyr	Leu	Arg
	335		340		345
Asp	Cys	Ile	Phe	Asp	Lys
	350		355		360
Ser	Ala	Ile	Met	Thr	Gly
	365		370		375
Asp	Arg	Asp	Thr	Arg	Arg
	380		385		390
Ser	Ser	Lys	Pro	Arg	Ala
	395		400		405
Gly	Gly	Lys	Arg	Arg	Lys
	410		415		420
Phe	Asn	Lys	Lys	Ile	Leu
	425		430		435
Val	Ile	Ala	Val	Ala	Ala

450

```
<220>  
<221> misc_feature  
<223> Incyte Clone Number: 132240
```

<400>	32						
cttttcctgg	aatttctata	atggaaagtc	cattagaaag	tcagccctta	gattcagata		60
gaagcatcaa	agaatcctct	tttgaagaat	caaataattga	agatccactt	attgtaacac		120
cagattgcc	agaaaagacc	tcaccaaag	gtgtcgagaa	ccctgctgta	caagagagta		180
acaaaaaaat	gttaggtcct	cctttggagg	tgtgaaaaac	gttagcctct	aaaagaaatg		240
ctgttgcttt	tcgaagtttt	aacagctata	ttaattgcac	caataactca	gaacctcca		300
gaatgaacat	gactttctta	gatgcaatgg	atatttcgtg	tgctcacagt	ggctcatatc		360
ccatggctat	aaccctact	caaaaaagaa	gatcctgtat	gccacatcag	accccaaatac		420
agatcaagtc	gggaactcca	taccgaactc	cgaagagtgt	gagaagaggg	gtggcccccg		480
ttgatgatgg	gcgaattcta	ggaaccccag	actaccttgc	acctgagctg	ttactaggca		540
gggcccattg	tcctgcggtg	gactgggtgg	cacttgaggt	ttgcttgttt	gaatttctaa		600
caggaattcc	ccctttcaat	gatgaaacac	cacaacaagt	attccagaat	attctgaaaa		660
gagatatccc	ttggccagaa	ggtgaagaaa	agttatctga	taatgctcaa	agtgcatgat		720
aaataacttt	aaccattgat	gatacaaaag	gagctggaat	gaaagagcta	aaacgtcatc		780
ctctcttcag	tgatgtggac	tgggaaaatc	tgcagcatca	gactatgcct	ttcatcccc		840
agccagatga	tgaaacagat	acctcctatt	ttgaagccag	gaatactgct	cagcacctga		900
ctgtatctgg	atttagtctg	tagcacaaaa	attttccttt	tagtctagcc	ttgtgttata		960
gaatgaactt	gcataattat	atactcctta	atactagatt	gatctaaggg	ggaaagatca		1020
ttattttaacc	tagttcaatg	tgcttttaat	gtacgttaca	gctttcacag	agttaaaagg		1080
ctgaaaggaa	tatagtcagt	aattttatctt	aacctcaaaa	ctgtatataa	atcttcaaag		1140
cttttttcat	ttattttatt	gtttttattgc	actttatgaa	aactgaagca	tcaataaaaat		1200
tagaqqacac	taaaaaaaaaa	a					1221

```
<220>
<221> misc_feature
<223> Incyte Clone Number: 2180116
```

400							
tgccaggct	gggtccagca	gcgcgatggc	agctcagcgg	ctgggcaagc	gcgtgctgag	60	
caagctgcag	tctccatcgc	gggcccgcgg	gccagggggc	agtcccgggg	ggatgcagaa	120	
gcggcacgcg	cgcgtcaccg	tcaagtatga	ccggcgggag	ctgcagcggc	ggctggacgt	180	
ggagaagtgg	atcgacgggc	gcctggaggga	gctgtaccgc	ggcatggagg	cagacatgcc	240	
cgatgagatc	aacattgatg	aattgttgga	gttagagagt	gaagaggaga	gaagccggaa	300	
aatccaggga	ctcctgaagt	catgtgggaa	acctgtcgag	gacttcattc	aggagctgct	360	
ggcaaagctt	caagggcttc	acaggcagcc	cggcctccgc	cagccaagcc	cctcccacga	420	
cggcagcctc	agccccctcc	aggaccgggc	cgggactgct	cacccctgac	cctcttgcaac	480	
tctccctgcc	ccccqqacgc	gcgccagctt	qcttqtqtat	aaqtqtatt	taatggattc	540	

tt

542

<210> 34  
 <211> 2778  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 2548, 2557, 2645, 2722, 2557, 2762, 2765  
 <223> Incyte Clone Number: 2197671

<400> 34  
 cgcggtatcgt cgcgggcccg ccgtcccgtc ccaggaagtg gccgtcctga gcgccatggc 60  
 tcaactcccc gtgcagtcgg gcctgcccgg catgcagaac ctaaaggcag acccagaaga 120  
 gctttttaca aaactagaga aaattgggaa gggctccttt ggagagggtgt tcaaaggcat 180  
 tgacaatcgg actcagaaag tgggtgccat aaagatcatt gatctggaag aagctgaaga 240  
 tgagatagag gacattcaac aagaaatcac agtgctgagt cagtgtgaca gtccatattgt 300  
 aaccaaatat tatggatcct atctgaagga taaaaatta tgataataa tggaatatct 360  
 tgggtggaggc tccgcactag atctattaga acctggccga ttagatgaaa ccagatcgc 420  
 tactatatta agagaaatag tgaaaggact cgattatctc cattcggaga agaaaatcca 480  
 cagagacatt aaagcggcca acgtcctgct gtctgagcat ggcgagggtga agctggcgga 540  
 ctttggcgtg gctggccagc tgacagacac ccagatcaaa aggaacacct tcgtgggcac 600  
 cccattctgg atggcaccgg aggtcatcaa acagtcggcc tatgactcga aggcagacat 660  
 ctggtccctg ggcataacag ctattgaact tgcaagaggg gaaccacctc attccgagct 720  
 gcaccccatg aaagtattat tcctcattcc aaagaacaac ccaccgacgt tgggaaggaaa 780  
 ctacagtaaa cccctcaagg agtttgtgga ggcctgtttg aataaggagc cgagcttttag 840  
 acccactgct aaggagttat tgaagcaca gtttatacta cgcaatgcaa agaaaacttc 900  
 ctacttgacc gagctcatcg acaggtacaa gagatggaag gccgagcaga gccatgacga 960  
 ctcgagctcc gaggattccg acgcggaaac agatggccaa gcctcggggg gcagtgttc 1020  
 tggggactgg atcttcacaa tccgagaaaa agatcccaag aatctcgaga atggagctct 1080  
 tcagccatcg gacttggaca gaaataagat gaaagacatc ccaaaggaggc ctttctctca 1140  
 gtgtttatct acaattattt ctctctgtt tgcagagttg aaggagaaga gccaggcgtg 1200  
 cggagggaac ttgggggtcca ttgaagagct gcgaggggcc atctacctag cggaggaggc 1260  
 gtgccctggc atctccgaca ccatggtggc ccagctcgtg cagcggctcc agagatactc 1320  
 tctaagtggt ggaggaactt catcccactg aaattccttt ggcatttggg gttttgtttt 1380  
 tcttttttct cttcttcac cctctccttt tttaaaagtc aacgagagcc ttcgctgact 1440  
 ccacogaaga ggtgcgccac tgggagccac ccagtgcca ggcgcccgtc caggagacaca 1500  
 cacagtcttc actgtgctgc agccagatga agtctctcag atgggtgggg agggtcagct 1560  
 ccttccagcg atcattttat tttattttat tacttttgtt ttttaattta accatagtgc 1620  
 acatattcca ggaaagtgtc tttaaaaaca aaaacaaacc ctgaaatgta tatttgggat 1680  
 tatgataagg caactaaaga catgaaacct caggtatcct gctttaagtt gataactccc 1740  
 tctgggagct ggagaatcgc tctggtggat ggggtgtacag atttgtatat aatgtcattt 1800  
 ttacggaaac ctttcggcg tgacataagga atcactgtgt acaaaactggc caagtgcctc 1860  
 tgtagataac gtcagtggag taaatattcg acaggccata acttgagctc attgccttgc 1920  
 ctttattaca tgtacatttt gaattctgtg accagtgatt tgggttttat tttgtatttg 1980  
 cagggtttgt cattaataat taatgcccct ctcttacaga acactcctat ttgtacctca 2040  
 acaaatgcaa attttcccg tttgcccctac gcccttttg gtacacctag aggttgattt 2100  
 cctttttcat cgatggtact atttcttagt gttttaaatt ggaacatatc ttgcctcatg 2160  
 aagctttaaa ttataatttt cagtttctcc ccatgaagcg ctctcgtctg acatttgttt 2220  
 ggaatcgtgc cactgctggt gtgcgccaga tgtaccgtcc tttccaatac gattttctgt 2280  
 tgcaccttgt agtggattct gcataatcct tttccacct aaaaatgtct gaatgcttac 2340  
 acaataaat tttataacac gcttattttg catactcctt gaaatgtgac tcttcagagg 2400  
 acagggtacc tgctgtgtat gtgtggccgt gcgtgtgtac tcgtggctgt gtgtgtgtga 2460

```

tgagacactt tggaaagactc cagggagaag ttcccagggc tggagctgcc gagtggccag 2520
gtcagcgccc tgggctgctt gcgcaatngc tcaccngat gatgcattgg aggttgctga 2580
cctgtgcgat tgctgtagcg gttgccaggg accttaaggg gttattttgc ttccctggga 2640
gggncctat gtttctaggg aagcagccat gtgtctaatt ttctgggttt gctgtgggga 2700
cctgattggg ggagggggaa anctttgggg ttcttgaggt gggaggggtc gtgccancaa 2760
tnttncctgg taaaaaag                                     2778

```

&lt;210&gt; 35

&lt;211&gt; 1424

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2594943

&lt;400&gt; 35

```

ggctcagcct ccgacccagg tggctctggag cctgccggga gagtgggtggc atctgagagg 60
ctggctcgtgg actgtggttg ggggaggtgg gagctgtttt aaccgtgtgc cccctctcct 120
gtgccggcgt gggcatcccc cggggcagtg gaacgcgggc gtcctccag cttccgagtc 180
cagccagcct gggcgcgggg cgccgcccc gagacacccg aggagtcctg tcctccctgg 240
ttacgtggac tgtggagctg gtctcttggt gtcagcgcc gtgcggaggt tgaagcgtag 300
ctgcggaggt cgcaccaggg cgtgaggagg aggaggaagg gcatgagccg agcttgagga 360
atccgtgctc caaactctac actcaagggt ggcccttggg tagggtgaag atcccctgtc 420
tttatactag ttccacacct tgggtgtgggt tactgggtgc aggatgaact gtcgctcgga 480
ggtgctggag gtgtcggtgg aggggcggca ggtggaggag gccatgctgg ctgtgctgca 540
cacggtgctt ctgcaccgca gcacaggcaa gttccactac aagaaggagg gcacctactc 600
cattggcacc gtgggcaccc aggatgttga ctgtgacttc atcgacttca cttatgtgctg 660
tgtctcttct gaggaactgg atcgtgccct gcgcaagggt gtgggggagt tcaaggatgc 720
actgcgcaac tctggtggcg atgggctggg gcagatgtcc ttggagtctt accagaagaa 780
gaagtctcgc tggccattct cagacgagtg catcccatgg gaagtgtgga cgggtcaagg 840
gcatgtggta gccctggcca cggagcagga gcggcagatc tgccgggaga aggtgggtga 900
gaaactctgc gagaagatca tcaacatcgt ggaggtgatg aatcggcatg agtacttgcc 960
caagatgccc acacagtcgg aggtggataa cgtgtttgac acaggcttgc gggacgtgca 1020
gccctacctg tacaagatct ccttcagat cactgatgcc ctgggcacct cagtcaccac 1080
caccatgcgc aggtcatca aagacaccct tgccctctga gcgtcgctgg atctctggga 1140
gtccttgat ggctcccaga ccttggtttt tgggaattgc acttttgggc ctttgggctc 1200
tggaacctgc tctgggtcat tgggtgagact tgggaagggc agccccgct ggcttcttgg 1260
ttttgtggtt gccagcctca ggtcatcctt ttaatctttg ctgatgggtc agtcctgcct 1320
ctactgtctc tccatagccc tgggtgggtc cccctctttt ctccactgta cagaagagcc 1380
accactggga tggggaataa agttgagaac atgaaaaaaa aaaa                                     1424

```

&lt;210&gt; 36

&lt;211&gt; 1839

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 1513871

&lt;400&gt; 36

```

cctcctctc gcgcagctca ggttgagct tctctgggga actgctcacc tttccggagc 60
aggggaagct gccccgtgcc cgggagggag cgggcgcacc gcggcccca ggacacgcgc 120
tgaccggct gccagtcct tcatgatcat gaacaagatg aagaacttta agcgcgcttt 180
ctccctgtca gtgccccgca ctgagaccat tgaagaatcc ttggctgaat tcacggagca 240

```



```

attcaaccag ctccacaacc ggcggaatga gaacttgag ctcggtcctc ttggcagaga 300
ccccccgcag gagtgcagca ccttctcccc aacagacagc ggggaggagc cggggcagct 360
ctcccctggc gtgcagttcc agcggcgga gaaccagcg cgcttctcca tggaggacgt 420
cagcaagagg ctctctctgc ccatggatat ccgcctgccc caggaattcc tacagaagct 480
acagatggag agcccagatc tgcccaagcc gctcagccgc atgtcccgcg gggcctccct 540
gtcagacatt ggctttggga aactggaaac atacgtgaaa ctggacaaac tgggagaggg 600
cacctatgcc acagtcttca aagggcgag caaactgagc gagaaccttg tggccctgaa 660
agagatccgg ctggagcacg aggagggagc gccctgcact gccatccgag aggtgtctct 720
gctgaagaac ctgaagcacg ccaatattgt gacctgcat gacctcatcc acacagatcg 780
gtccctcacc ctggtgtttg agtacctgga cagtgcctg aagcagtatc tggaccactg 840
tgggaacctc atgagcatgc acaacgtcaa gattttcatg ttccagctgc tccggggcct 900
cgctactgt caccaccgca agatcctgca ccgggacctg aagccccaga acctgctcat 960
caacgagagg ggggagctga agctggccga ctttggactg gccagggcca agtcagtgc 1020
caciaaagact tactccaatg aggtggtgac cctgtgttac agggcccccg atgtgctgct 1080
gggatccaca gactactcca cccccattga tatgtggggc gtgggctgca tccactacga 1140
gatggccaca gggaggcccc tcttcccggg ctccacagtc aaggaggagc tgcacctcat 1200
ctttcgctc ctccgggacc ccacagaaga gacgtggccc ggcgtgaccg ctttctctga 1260
gttcgcgacc tacagcttcc cctgctacct ccgcagccg ctcatcaacc acgcgcccag 1320
gttgatagc gatggcatcc acctcctgag cagcctgctc ctgtatgaat ccaagagtcg 1380
catgtcagca gaggtgccc tgagtcactc ctacttccgg tctctgggag agcgtgtgca 1440
ccagcttgaa gacctgect ccatcttctc cctgaaggag atccagctcc agaaggacc 1500
aggctaccga ggcttggcct tccagcagcc aggacgagg aagaacaggc ggcagagcat 1560
cttctgagcc acgcccacct tgcgtgggac aagggacaag agatcacatg gagcacaat 1620
tcgggtagga tggagcctgt gtggccctcg gagactgaa gaacgagggc tgacagcagc 1680
ctggaagacc gcttggcagg cttttggcca agtggttttc tttgtggtt cgatctgctg 1740
ccagtagttt cagtggatc aacgtgcttt aggagttggg tgggaaagtc ttgctagagg 1800
gtttaggggg aggtttctac cgttgactcg gtttagggc 1839

```

&lt;210&gt; 37

&lt;211&gt; 2024

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 156108

&lt;400&gt; 37

```

gtcagctctg gttcggagaa gcagcggctg gcgtgggcca tccggggaat gggcgccctc 60
gtgacctagt gttcgggggc aaaaagggtc ttgcccgcct cgctcgtgca gggcgctatc 120
tgggcgcctg agcgcggcgt gggagccttg ggagccgccc cagcaggggg cacaccgga 180
accggcctga gcgcccggga ccatgaacgg ggaggccatc tgcagcgccc tgcccaccat 240
tccctaccac aaactcgccg acctgcgcta cctgagccgc ggcgcctctg gcaactgtgtc 300
gtccgcccgc cagcagact ggcgcgtcca ggtggcctg aagcacctgc acatccacac 360
tccgctgctc gacagtgaaa gaaaggatgt cttaagagaa gctgaaattt tacacaaagc 420
tagatttagt tacattcttc caattttggg aatttgcaat gagcctgaat ttttgggaat 480
agttactgaa tacatgcaa atggatcatt aaatgaactc ctacatagga aaactgaata 540
tctgatgtt gcttggccat tgagatttgc catectgcat gaaattgccc ttggtgtaaa 600
ttacctgcac aatatgactc ctcccttact tcatcatgac ttgaagactc agaatactt 660
attggacaat gaatttcatg ttaagattgc agattttggt ttatcaaagt ggcgcatgat 720
gtccctctca cagtacgaa gtagcaaatc tgcaccagaa ggagggacaa ttatctatat 780
gccacctgaa aactatgaac ctggacaaaa atcaagggcc agtatcaagc acgatataata 840
tagctatgca gttatcacat gggagtggtt atccagaaaa cagccttttg aagatgtcac 900
caatcctttg cagataatgt atagtgtgtc acaaggacat cgacctgtta ttaatgaaga 960
aagtttgcca tatgatatac ctaccagagc acgtatgatc tctctaataa aaagtggatg 1020

```

```

ggcacaaaat ccagatgaaa gaccatcttt cttaaaatgt ttaatagaac ttgaaccagt 1080
tttgagaaca tttgaagaga taacttttct tgaagctgtt attcagctaa agaaaacaaa 1140
gttacagagt gtttcaagtg ccattcacct atgtgacaag aagaaaatgg aattatctct 1200
gaacatacct gtaaatcatg gtccacaaga ggaatcatgt ggatcctctc agctccatga 1260
aaatagtggt tctcctgaaa cttcaaggtc cctgccagct cctcaagaca atgatttttt 1320
atctagaaaa gctcaagact gttattttat gaagctgcat cactgtcctg gaaatcacag 1380
ttgggatagc accatttctg gatctcaaag ggctgcattc tgtgatcaca agaccactcc 1440
atgctcttca gcaataataa atccactctc aactgcagga aactcagaac gtctgcagcc 1500
tggtatagcc cagcagtggg tccagagcaa aagggaagac attgtgaacc aaatgacaga 1560
agcctgcctt aaccagtcgc tagatgccct tctgtccagg gacttgatca tgaaagagga 1620
ctatgaactt gttagtacca agcctacaag gacctcaaaa gtcagacaat tactagacac 1680
tactgacatc caaggagaag aatttgccaa agttatagta caaaaattga aagataacaa 1740
acaaatgggt cttcagcctt acccggaat acttggtgtt tctagatcac catctttaa 1800
tttacttcaa aataaaaagca tgtaatgac tgtttttcaa gaagaaatgt gtttcataaa 1860
aggatattta tatctctgtt gctttgactt tttttatata aaatccgtga gtattaaagc 1920
tttattgaag gttctttggg taaatattag tctccctcca tgacactgca gtattttttt 1980
taattaatac aagtaaaaag tttgaatttt gctacataaa aaaa 2024

```

&lt;210&gt; 38

&lt;211&gt; 1861

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2883243

&lt;400&gt; 38

```

gcttcttagt gaggttggca ttatgttaag gctggtatgg aagacaactg atgaagcagg 60
agtggctctg tgacattttt ctgacttgat tggctggggc gtgtgatgta atagggttca 120
gtgcagcccc ttataggttt taaaatgaat tccaagacac cattacaaag aaagccggac 180
tcttttctta taactgagct cagccaagga aactcttgca caaatgtaca atactgtttg 240
gaatatggaa gacctggatt tagaatatgc caagacagat ataaattgtg gcacagactt 300
gatgttttat atagaaatgg acccaccagc actgcctcct aaaccaccaa aacctactac 360
tgtagccaac aacggtatga ataacaatat gtccttaca gatgctgaat ggtactgggg 420
agatatctcg agggaagaag tgaatgaaaa acttcgagat acagcagacg ggaccttttt 480
ggtacgagat gcgtctacta aaatgcatgg tgattatact cttacactaa ggaagggggg 540
aaataacaaa ttaatcaaaa tatttcatcg agatgggaaa tatggcttct ctgacctatt 600
aaccttcagt tctgtggttg aattaataaa ccactaccgg aatgaatctc tagctcagta 660
taatcccaaa ttggatgtga aattacttta tccagtatcc aaataccaac aggatcaagt 720
tgtcaaaaga gataatattg aagctgtagg gaaaaaatta catgaatata acactcagtt 780
tcaagaaaaa agtcgagaat atgatagatt atatgaagaa tataccgca catcccagga 840
aatccaaatg aaaaggacag ctattgaagc atttaatgaa accataaaaa tatttgaaga 900
acagtgccag acccaagagc ggtacagcaa agaatacata gaaaagttaa aacgtgaagg 960
caatgagaaa gaaatacaaa ggattatgca taattatgat aagttgaagt ctggaatcag 1020
tgaaattatt gacagtagaa gaagattgga agaagacttg aagaagcagg cagctgagta 1080
tcgagaaatt gacaaacgta tgaacagcat taaaccagac cttatccagc tgagaaagac 1140
gagagaccaa tacttgatgt ggttgactca aaaagggtgt cggcaaaaga agttgaacga 1200
gtggttgggc aatgaaaaca ctgaagacca atattcactg gtggaagatg atgaagattt 1260
gccccatcat gatgagaaga catggaatgt tggaagcagc aaccgaaaca aagctgaaaa 1320
cctgttgcca gggaaagcag atggcacttt tcttgtccgg gagagcagta aacaggggctg 1380
ctatgcctgc tctgtagtgg tggacggcga agtaaagcat tgtgtcataa acaaaacagc 1440
aactggctat ggctttgccg agccctataa cttgtacagc tctctgaaag aactggtgct 1500
acattaccaa cacacctccc ttgtgcagca caacgactcc ctcaatgtca cactagccta 1560
cccagtatat gcacagcaga ggcgatgaag cgcttactct ttgatccttc tcctgaagtt 1620

```

```

cagccaccct gaggcctctg gaaagcaaag ggctcctctc cagtctgac tgtgaattga 1680
gctgcagaaa cgaagccaac tttttttgga tgggactagt gctttctttc acaaaaaaga 1740
agtaggggaa gacatgcagc ctaaggctgt atgatgacca cagttccta agctggagt 1800
cttatccctt ctttttcttt ttttctttgg tttaatttaa agccacaacc acatacaaca 1860
c

```

```

<210> 39
<211> 2045
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte Clone Number: 3173355

```

```

<400> 39
cttggtcgga acctgagacg gattcgctcc caaatgatgc tccagtggca ggagcaactc 60
aagttcatca ttgtcctgag agagaggagc agcgcggttc tcggccggga cagcagaacg 120
ccaggggacc ctccacctgg cgcgcggggg caggggcttt gattgtcctg gggtcgcgga 180
gaccgcgcgc cctgcctctg acgcccggcg gcaacctttg cagtgcggtt gggtgctgcg 240
atcgcccggc gggtcctctg cgaaggctcg gctgcttctg tccacctctt acacttcttc 300
atttatcggt ggatcatttc gagagtcctg cttgtaaatg tttggcactt tgctacttta 360
ttgcttcttt ctggcgacag ttccagcact cgccgagacc ggccgagaaa ggagctgag 420
cccggagaag agcgaaatat ggggacccgg gctaaaagca gacgtcgtcc tccccgccc 480
ctatttctat attcaggcag tggatacatc agggataaaa ttcacatctt ctccaggcga 540
aaaggtcttc caggtgaaag tctcagcacc agaggagcaa ttcactagag ttggagtcca 600
ggttttagac cgaaaagatg ggtccttcat agtaagatac agaattgtat caagctacaa 660
aaatctgaag gtggaaatta aattccaagg gcaacatgtg gccaaatccc catatatatt 720
aaaaagggcg gtttaccatg agaactgtga ctgtcctctg caagatagtg cagcctggct 780
acgggagatg aactgccctg aaaccattgc tcagattcag agagatctgg cacatttccc 840
tgctgtggat ccagaaaaga ttgcagtaga aatcccaaaa agatttggac agaggcagag 900
cctatgtcac tacaccttaa aggataacaa ggttttatatc aagactcatg gtgaacatgt 960
aggttttaga attttcatgg atgccatact actttctttg actagaaagg tgaagatgcc 1020
agatgtggag ctctttgtta atttgggaga ctggcctttg gaaaaaaga aatccaattc 1080
aaacatccat ccgatctttt cctggtgtgg ctccacagat tccaaggata tcgtgatgcc 1140
tacgtacgat ttgactgatt ctgttctgga aaccatgggc cgggtaagtc tggatatgat 1200
gtccgtgcaa gctaacacgg gtccctcctg ggaaagcaaa aattccactg ccgtctggag 1260
agggcgagac agccgcaaag agagactcga gctggttaa ctcagtagaa aacaccaga 1320
actcatagac gctgctttca ccaacttttt cttctttaaa cacgatgaaa acctgtatgg 1380
tcccattgtg aaacatatatt cattttttga tttcttcaag cataagtatc aaataaatat 1440
cgatggcact gtagcagctt atcgccctgc atatttgcta gttggtgaca gtgtgtgct 1500
gaagcaggat tccatctact atgaacattt ttacaatgag ctgcagccct ggaaacacta 1560
cattccagtt aagagcaacc tgagcgatct gctagaaaaa cttaaatggg cgaaagatca 1620
cgatgaagag gccaaaaaga tagcaaaagc aggacaagaa tttgcaagaa ataattctcat 1680
gggcgatgac atattctgtt attatttcaa acttttccag gaatatgcca atttacaagt 1740
gagtgaagccc caaatccgag agggcatgaa aagggtagaa ccacagactg aggacgacct 1800
cttcccttgt acttgccata ggaaaaagac caaagatgaa ctctgatatg caaaaataact 1860
tctattagaa taatggtgct ctgaagactc ttcttaacta aaaagaagaa tttttttaag 1920
tattaattcc atggacaata taaaatctgt gtgattgtt gcagtatgaa gacacatttc 1980
tacttatgca gtattctcat gactgtactt taaagtacat ttttagaatt ttataataaa 2040
accac

```

```

<210> 40

```

<211> 1260

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 5116906

<400> 40

```
cgatattttt ctttcttagt ttcccatctt atattgtttt gtcaaataca ctgtgactca 60
ttaacatctc ttttccctag gttttgctgg cacacctgga tatctttctc cagaagtgtt 120
acgtaaagat ctttatggaa agccagtgga tatgtgggca tgggtgtca ttctctatat 180
tctactttgt gggatccac cttctggga tgaagacca cacagactct atcagcagat 240
caaggctgga gcttatgatt ttccatcacc agaattgggac acggtgactc ctgaagccaa 300
agacctatc aataaaatgc ttactatcaa ccctgccaaa cgcatacacag cctcagaggc 360
actgaagcac ccatggatct gtcaacgttc tactgttgct tccatgatgc acagacagga 420
gactgtagac tgcttgaaga aatttaatgc tagaagaaaa cttaaaggtg ccatcttgac 480
aactatgctg gctacaagga atttctcagc agccaagagt ttgttgaaga aaccagatgg 540
agtaaaggag tcaactgaga gttcaaatac aacaattgag gatgaagatg tgaaagcacg 600
aaagcaagag attatcaaag tcaactgaac actgatcgaa gctatcaaca atggggactt 660
tgaagcctac acaaaaatct gtgaccagg cttactgct ttgaaacctg aagctttggg 720
taatttagtg gaaggatgg atttccaccg atttacttt gaaaatgctt tgtccaaaag 780
caataaacca atccacacta ttattctaaa ccctcatgta catctggtag gggatgatgc 840
cgctgcata gcatatatta ggctcacaca gtacatggat ggcatggaa tgccaaagac 900
aatgcagtca gaagagactc gtgtgtggca ccgccgggac ggaaagtggc agaattgtca 960
ttttcatcgc tcggggtcac caacagtacc catcaactaa atttcaacag tgccacttct 1020
gcattctctg ttctcaaggc acctggatgg tgacctggg ccgtcctctc ctctcttca 1080
tgcatgtttc tgagtgcagt aagttgtgaa ggctctacat gtaatgcata tgtgatgcat 1140
catcttatca tatattcctt cctatacatt gtttactct caactacggg gatgttccac 1200
acaaacttaa attactgttg gcaaaacaat agggggagat tagacaaaaa aaaaaaaaaa 1260
```

<210> 41

<211> 2059

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 940589

<400> 41

```
aaaccataga aacgctaatt aaagcagaca tcaaaatctg gatccttaca ggggacaagc 60
aagaaactgc cattaacatc ggacactcct gcaaaactgt gaagaagaac atgggaatga 120
ttgttataaa tgaaggctct cttgattctt tctctaatac acagaattct aggaaggagg 180
ctgttctttt agccaaaatg aaacacccta atattgttgc cttcaaagaa tcatttgaag 240
ctgaaggaca cttgtatatt gtgatggaat actgtgatgg aggggatcta atgcaaaaga 300
ttaaacagca gaaaggaaag ttatttcctg aagacatgat acttaattgg tttacccaaa 360
tgtgccttgg agtaaatcac attcacaaga aacgtgtgct acacagagat atcaagtcca 420
agaatatctt cctcactcag aatggaaaag tgaaattggg agactttgga tctgcccgtc 480
ttctctccaa tccgatggca tttgcttgta cctatgtggg aactccttat tatgtgcctc 540
cagaaatttg ggaacacctg cttataaaca ataaaagtga catctggtcc ttgggttgca 600
tcctgtatga actctgtacc cttaagcatc catttcaggc aaatagtgtg aaaaatctta 660
tcctcaaagt atgtcaaggg tgcacagtc cactgccgtc tcattactcc tatgaacttc 720
agttcctagt caagcagatg tttaaaagga atcctcaca tcgcccctcg gctacaacgc 780
ttctctctcg aggcacgta gctcggcttg tccagaagtg cttaccccc gagatcatca 840
```

```

tggaatatgg tgaggaagta ttagaagaaa taaaaaatc gaagcataac acaccaagaa 900
aaaaaacaaa cccagcaga atcaggatag ctttgggaaa tgaagcaagc acagtgcag 960
aggaagaaca agatagaaag ggtagccata ctgatttga aagcattaat gaaaatttag 1020
ttgaaagtgc attgagaaga gtaaacagag aagaaaaagg taataagtca gtccatctga 1080
ggaaagccag ttcaccaa atccatagac gacagtggga gaaaaatgta cccaatacag 1140
ctcttacagc tttggaaaat gcatccatac tcacctccag tttaacagca gaggacgata 1200
gaggtggttc tgtaataaag tacagcaaaa atactactcg taagcagtgg ctcaaagaga 1260
ccccggacac tttgttgaac atccttaaga atgctgatct cagcttggct tttcaaact 1320
acacaatata tagaccaggt tcagaagggt tcttgaaagg cccctgtct gaagaaacag 1380
aagcatcgga cagtgttgat ggaggtcacg attctgtcat tttggatcca gagcgacttg 1440
agcctgggct agatgaggag gacacggact ttgaggagga agatgacaac cccgactggg 1500
tgtcagagct gaagaagcga gctggatggc aaggcctgtg cgacagataa tgcctgagga 1560
aatgttcctg agtcacgctg aggagagcct tcactcagga gttcatgctg agatgatcat 1620
gagttcatgc gacgtatatt ttcctttgga aacagaatga agcagaggaa actcttaata 1680
cttaaaatcg ttcttgatta gtatcgtgag tttgaaaagt ctagaactcc tgtaagtgtt 1740
tgaactcaag ggagaaggta tagtggaatg agtgtgagca tcgggcttg cagtcaccata 1800
gaacagaaat gggatgctag cgtgccacta cctacttgtg tgattgtggg aaattactta 1860
acctcttcaa gccccaat tctcaacat aaaatgaaga taataatgcc tacctcagag 1920
ggatgctgac cacagacctt tatagcagcc cgtatgatat tattcacatt atgatatgtg 1980
tttattatta tgtgactctt tttacatttc ctaaagggtt gagaattaaa tatatttaat 2040
tatgaaaaaa aaaaaaaaaa 2059

```

&lt;210&gt; 42

&lt;211&gt; 1023

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 304421

&lt;400&gt; 42

```

gaggcagagg ggtgggcggg ctggcccatg gctgagacct ctctcccaga gctgggggga 60
gaggacaaaag ccacgccttg cccagcctc ctggagctgg aggagctcct gcgggcaggg 120
aagtcttctt gcagccgtgt ggacgaagtt tggcccaacc ttttcatagg agatgcgatg 180
gactcactgc agaagcagga cctccggagg cccaagatcc atggggcagt ccaggcatct 240
ccctaccagc cgccacatt ggcttcgctg cagcgcttgc tgtgggtccg tcaggctgcc 300
acactgaacc atatcgatga ggtctggccc agcctcttcc tgggagatgc gtacgcagcc 360
cgggacaaga gcaagctgat ccagctggga atcaccacg ttgtgaatgc cgctgcaggc 420
aagttccagg tggacacagg tgccaaatc taccgtggaa tgtccctgga gtactatggc 480
atcgaggcgg atgacaaccc cttcttcgac ctcagtgtct actttctgcc tgttgcctga 540
tacatccgag ctgccctcag tgttcccaa ggccgcgtgc tggtagactg tgccatgggg 600
gtaagccgct ctgccacact tgtcctggcc ttctcatga tctatgagaa catgacgctg 660
gtagaggcca tccagacggt gcaggccac cgcaatatct gccctaactc aggtctcctc 720
cggcagctcc aggttctgga caaccgactg gggcgggaga cggggcggt ctgatctggc 780
aggcagccag gatccctgac ccttgcccaa accccaccag cctggccctg ggaacagcag 840
gctctgctgt ttctagtgc cctgagatgt aaacagcaag tgggggctga ggcagaggca 900
gggtagctg ggtggtgacc tcttagcggg tggatttccc tgacccaatt cagagattct 960
ttatgcaaaa gtgagttcag tccatctcta taataaaata ttcctcgtca taaaaaaaaa 1020
aaa

```

&lt;210&gt; 43

&lt;211&gt; 4416

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 1213802

&lt;400&gt; 43

```

gaaatTTTTt tctgcctcat tattattaat tcatggattg agtgttggtt cgacctacag 60
gcgtaataga ttggaactca gtgaagacac agatgttcct gttcagagca accagctaatt 120
gattacagtt taaagacaat ttctgtgata aagttgtcat ttggaagatt aaacccattt 180
cacgaggact tggagcctgg tccttgcttt gaggaagcag tggcttggtt caagaagcca 240
cttctgactt aagaatctac ccagcatgcc taatcaagga gaagactgct attttttttt 300
ctattccaca tgtaccaaag gcgacagctg ccattccgt cactgtgaag ctgcaatagg 360
aatgaaact gtttgcacat tatggcaaga agggcgctgt tttcgacagg tgtgcagggt 420
tcggcacatg gagattgata aaaaacgcag tgaaattcct tgttattggg aaaatcagcc 480
aacaggatgt caaaaattaa actgcgcttt ccatcacaat agaggacgat atgttgatgg 540
ccttttccta cctccgagca aaactgtgtt gccactgtg cctgagtcac cagaagagga 600
agtgaaggct agccaacttt cagttcagca gaacaaattg tctgtccagt ccaatccttc 660
ccctcagctg cggagcggtta tgaaagtaga aagttccgaa aatgttccta gccccacgca 720
tccaccagtt gtaattaatg ctgcagatga tgatgaagat gatgatgatc agttttctga 780
ggaaggtgat gaaacaaaa cacctaccct gcaaccaact cctgaagttc acaatggatt 840
acgagtgact tctgtccgga aacctgcagt caatataaag caaggtgaat gtttgaattt 900
tggaataaaaa actcttgagg aaattaagtc aaagaaaatg aaggaaaaat ctaagaagca 960
aggtgagggg tcttcaggag tttccagtct tttactccac cctgagcccg ttccagggtc 1020
tgaaaaagaa aatgtcagga ctgtggtgag gacagtaact ctctccacca aacaaggaga 1080
agaacccttg gttagattga gtcttactga gagactgggg aaacgaaaat tttcagcagg 1140
cggtgacagt gatcctccat taaagcgtag cctggcacag aggctaggga agaaagtga 1200
agctccagaa actaacattg acaaaacacc aaagaaagct caagtttcca agtctcttaa 1260
ggagcgatta ggcagtgcag ctgatccaga taatgaggat gcaacagata aagttaataa 1320
agttggtgag atccatgtga agacattaga agaaattctt cttgaaagag ccagtcagaa 1380
acgtggagaa ttgcaacta aactcaagac agaaggacct tcaaaaactg atgattctac 1440
ttcaggagca agaagctcct ccactatccg tatcaaaacc ttctctgagg tcctggctga 1500
aaaaaaacat cggcagcagg aagcagagag acaaaaaagc aaaaaggata caacttgcac 1560
caagctaaag attgatagt aaattaaaa aacagtagtt ttgccacca ttgttgccag 1620
cagaggacaa tcagaggagc ctgcaggtaa aacaaagtct atgcaggagg tgcacatcaa 1680
gacgctggaa gaaattaaac tggagaaggc actgaggggtg cagcagagct ctgagagcag 1740
caccagctcc cegtctcaac acgaggccac tcagggggca aggcggctgc tgcgaatcac 1800
caaaagaaca gggatgaaag aagagaagaa ccttcaggaa ggaaatgaag ttgattctca 1860
gagcagtatt agaacagaag ctaaaagggc ttcagggtgag accacaggag ttgacatcac 1920
taaaattcaa gtcaagagat gtgagaccat gagagagaag cacatgcaga aacagcagga 1980
gagggaaaaa tcagtcttga cacctcttcg gggagatgta gcctcttgca ataccaagt 2040
ggcagagaaa ccagtgtcca ctgctgtgcc aggaatcaca cggcacctga ccaagcggct 2100
tcccacaaag tcatcccaga aggtggaggt agaaacctca gggattggag actcattatt 2160
gaatgtgaaa tgtgcagcac agaccttgga aaaaaggggt aaagctaaac ccaaagtga 2220
cgtgaagcca tctgtggtta aagttgtgtc atcccccaa ttggcccaa aacgtaaggc 2280
agtggagatg cacgctgctg tcattgccgc tgtgaagcca ctcagctcca gcagtgtcct 2340
acaggaaacc ccagccaaaa aggcagctgt ggtgtgtgtc ccgcttgtct ctgaggacaa 2400
atcagtcact gtgctgaag cagaaaatcc tagagacagt cttgtgctgc ctccaacca 2460
gtcctcttca gattcctcac ccccgagggt gtctggccct tcctcatccc aatgagcat 2520
gaaaactcgc cgactcagct ctgcctcaac aggaaagccc cactctctg tggaggatga 2580
ttttgagaaa ctaatatggg agatttcagg aggcaaaattg gaagctgaga ttgacctgga 2640
tcctgggaaa gatgaagatg accttctgct tgagctatca gaaatgattg atagctgaag 2700
gtggtagtga ggacacttta aaaaaaaat cgccaaaaaa ctggacttag tttcatctat 2760
tgtaacattt acctgagatg atcatctctt tagtctagaa tttgcccmaa atcagaagta 2820
tacctctgaa ttatctgtat gtgtcctgga ttccttgggg tcagattttt aaagttactt 2880

```

```

tataaccatt ttgtccattt gatgccattg tttatcatct tttgagaaaa aagttctgtc 2940
atacccttct ctccacaaaa aagagactga gagggagatc aagtgaagg gtcgaagcga 3000
acttagtgac tccttgaggt gtttgtcagt tttggttttt ttcttctttg ttgtattctt 3060
tatgtattgt cttgatgtac ttaatattac ctgagtttga aatggatgaa gacagctgct 3120
accattaagg accaaatfff atgctaccac taaacaaaaa taccactca gtctgtgta 3180
aattgtatgt ctttttaaag gtatttaaag attcaactaa gctttaaaga gggctgagca 3240
gctcaggaag cctgtaatgt gggcataact ctttggacct gatcttgatg cttctgctgc 3300
tctgttagcc tctgaagagc aatatcta atattattac tgtaattttt taaaaggctt 3360
taaagtgcct caggggtccc ctgaaactaa ttttctattt ctgggattcc ctggattcat 3420
tatatgagat ggtgacatga ttagaggaat tcttttttag tatgaaaatt gtcccttttc 3480
ttcttcagta cttgcctcct tgctggcatt gaattaacac agggacaaaa tttggttaat 3540
tttttatttc taactctccc aacaaacccc tgttgcccag tatttgtttg gtggccttta 3600
accacctgag ggaaaaaatg agcttattca agctgccaat attatctat gggctgtagc 3660
agtacactga attgtactgt gccagggata ttgagatgct ctgggggtgt attgtatacc 3720
tgccagtttt cttcatttct gaattgagtt ttcttttctt gatgttggtt tccttcatat 3780
cacctcaagg tttagatttg tgaaggaata agcatgatgg aaataatagt cttgaaagga 3840
gatatgttgt atataatcag gaggaagagg aaggaaggac ttaccattt tgatattttg 3900
ctgtaggtgg ccagttttgt ttctcatagg gaaatctgac ccacctgtca tgttggtccc 3960
taaggaaactg ctgttgtaag cggctcatca agagttgaac ttcacgtagc cttgttgagg 4020
atatggaaaa ggaagaaagc cacaggactg cccattcagt cttgggaaga ttgggatgat 4080
tctgcacaag caaaaatgac tgaagtttat gtatagacac acctctacca atccatcttc 4140
agctgactga atgttgtatg atagcccttc tccaaagcag aggtagaatg ttcaggtttc 4200
accatggatt ttctacttat ttcgtttttg gaatcagctt acagattcca ggtccctttt 4260
gtatatattc tttattcttt tgctttttta aaaaataatt ttgtttcata tttaaagcac 4320
ttgtattagt caatgtttcg tgtccgcat tatttgaacc atttgccctt acagaaagag 4380
aaataactgt ttgtgtttta aataaaactg atgtag 4416

```

&lt;210&gt; 44

&lt;211&gt; 2068

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 1378134

&lt;400&gt; 44

```

gcagtcctac agtccgctga tgcgtcgccg ggccagcaac gctgcccgcg cagcccacac 60
gattggcggc agtaagcaca caatgaatga tcacctgcat gtcggcagcc acgctcacgg 120
acagatccag gttcgacagt tgtttgagga taacagtaac aagcggacag tgctcacgac 180
acaaccaa at gggcttacia cagtgggcaa aacgggcttg ccagtggtgc cagagcggca 240
gctggacagc attcatagac ggcaggggag ctccacctct ctaaagtcca tgggaaggcat 300
ggggaagggtg aaagccaccc ccatgacacc tgaacaagca atgaagcaat acatgcaaaa 360
actcacagcc ttcgaacacc atgagatttt cagctaccct gaaatatatt tcttggttct 420
aatgctaag aagcgccagg gcatgacagg tgggcccac aatggtggct atgatgatga 480
ccaggatca tatgtgcagg tgccccacga tcacgtggct tacaggtatg aggtcctcaa 540
ggtcattggg aaggggagct ttgggcaggt ggtcaaggcc tacgatcaca aagtcacaca 600
gcacgtggcc ctaaaatggg tgcggaatga gaagcgcttc caccggaag cagcggagga 660
gatccgaatc ctggaacacc tgcggaagca ggacaaggat aacacaatga atgtcatcca 720
tatgtgggag aatttcacct tccgcaacca catctgcatg acgtttgagc tgctgagcat 780
gaacctctat gagctcatca agaagaataa attccagggc ttcagtctgc ctttggttcg 840
caagtttgcc cactcgattc tgcagtgttc ggatgctttg cacaaaaaca gaataattca 900
ctgtgacctt aagcccgaga acattttgtt aaagcagcag ggtagaagcg gtattaaagt 960
aattgatttt ggctccagtt gttacgagca tcagcgtgtc tacacgtaca tccagtcgcg 1020
tttttaccgg gctccagaag tgatccttgg ggccaggat ggcatgccc ttgatattgt 1080

```

```

gagcctgggc tgcatttttag cagagctcct gacgggttac cccctcttgc ctggggaaga 1140
tgaaggggac cagctggcct gtatgattga actgttgggc atgccctcac agaaactgct 1200
ggatgcatcc aaacgagcca aaaattttgt gagctccaag gggtatcccc gttactgcac 1260
tgtcacgact ctctcagatg gctctgtggt cctaaacgga ggccgttccc ggagggggaa 1320
actgaggggc ccaccggaga gcagagagtg ggggaacgcg ctgaaggggt gtgatgatcc 1380
ccttttctct gacttcttaa aacagtgttt agagtgggat cctgcagtgc gcatgacccc 1440
aggccaggct ttgcggcacc cctggctgag gaggcggttg ccaaagcctc ccaccgggga 1500
gaaaacgtca gtgaaaagga taactgagag caccggtgct atcacatcta tatccaagtt 1560
acctccacct tctagctcag cttccaaact gaggactaat ttggcgca ga tgacagatgc 1620
caatgggaat attcagcaga ggacagtgtt gccaaaactt gttagctgag ctcacgtccc 1680
ctgatgctgg taacctgaaa gatacgacat tgctgagcct tactgggttg aaaaggagta 1740
gctcagacct gtttttattt gctcaataac tctactcatt tgtatctttt cagcacttaa 1800
ttttaatgta agaaaagtgt tcatTTTgtt ttataaaaat acatgaggac aatgctttta 1860
gtttttatac tttcagaaac tttttgtgtt ctaaaagtac aatgagcctt actgtattta 1920
gtgtggcaga ataataacat cagtggcagg ccactgatta cttcatgact gccacgcatt 1980
tacagattgg tgtcaaagac attcactatg tttttatggt tcatgttata tcctccccag 2040
ggtgacagcc ccttaaggcc ctcctttt 2068

```

<210> 45

<211> 1850

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 1490070

<400> 45

```

gggctgcctg cctgcctgcc tgcctgcctg gcccgcccg agctccagcc tgcctcttcc 60
actggccact gcctcccacc cagggtggc atccctgctc cctgccttgg gtcccagact 120
gtgtcctcca tcaccgcagg tcggtgaggg gctgggctgg acaccagggc ccgccctccc 180
atcactgagc tccactcctt cctcattttg ctgctgattc tagccccaaa caaaacaggt 240
tgagcttttt cctcccctca gaagctcctc tctggctcgt ggctgccttc tgagtgttgc 300
agacggcgcc ggccgggaag gggggcctgg gccagccctg ccaggactgg gacgctgctg 360
ctggcgccctg gccctccatc aggccagcct gtggcaggag agtgagcttt gccgcggcag 420
acgcctgagg atgatgcccc agctgcagtt caaagatgcc ttttggtgca gggacttcac 480
agcccacacg ggctacgagg tgctgctgca gcggttctg gatggcagga agatgtgcaa 540
agacatggtg gagctactgt ggcagagggc ccaggcggag gagcggtagc ggaaggagct 600
ggtgcagatc gcacggaagg cagggtggca gacggagatc aactccctga gggcctcctt 660
tgactccttg aagcagcaaa tggagaatgt gggcagctca cacatccagc tggccctgac 720
cctgcgtgag gagctgcgga gtctcgagga gtttcgtgag aggcagaagg agcagaggaa 780
gaagtatgag gccgtcatgg accgggtcca gaagagcaag ctgtcgctct acaagaaggc 840
catggagtcc aagaagacat acgagcagaa gtgccgggac gcggacgacg cggagcaggc 900
cttcgagcgc attagcgcca acggccacca gaagcaggtg gagaagagtc agaacaaagc 960
caggcagtg c aaggactcgg ccaccgaggc agagcgggta tacaggcaga gcattgcgca 1020
gctggagaag gtccgggctg agtgggagca ggagcaccgg accacctgtg aggcctttca 1080
gctgcaagag tttgaccggc tgaccattct ccgcaacgcc ctgtgggtgc acagcaacca 1140
gctctccatg cagtgtgtca aggatgatga gctctacgag gaagtgcggc tgacgctgga 1200
aggctgcagc atagacgccg acatcgacag tttcatccag gccaagagca cgggcacaga 1260
gcccccgct ccggtgccct accagaacta ttacgatcgg gaggtcaccc cgctgaccag 1320
cagccctggc atacagccgt cctgcggcat gataaagagg ttctctggac tgctgcacgg 1380
aagtcccaag accacttcgt tggcagcttc tgctgcgtcc acagagaccc tgacccccac 1440
ccccgagcgg aatgaggggtg tctacacagc catcgagtg caggagatac agggaaaccc 1500
ggcctcacca gcccaggagt accgggcgct ctacgattat acagcgca ga acccagatga 1560
gctggacctg tccgcgggag acatcctgga ggtgatcctg gaaggggagg atggctggtg 1620

```



```

gactgtggag aggaacgggc agcgtggcct cgctccctggt tcctacctgg agaagctttg 1680
aggaagggcc aggagccctc tcggacctgc cctggcagtg gagccagcag tgccccagc 1740
actgtcccca ccttgctagg gccagaacc aagcgtcccc cagccccgag agggagcctg 1800
tcgtctccca gggaataaag gagtgcgttc tgttctcaaa aaaaaaaaaa 1850

```

&lt;210&gt; 46

&lt;211&gt; 2534

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 1997814

&lt;400&gt; 46

```

gaagagggga tggagcaggg gctggaggag gaagaagagg tggatccccg gatccagggg 60
gaactggaga agttaaataca gtccacggat gatatacaaca gacgggagac tgaacttgag 120
gatgctcgctc agaagttccg ctctgttctg gttgaagcaa cggtgaaact ggatgaactg 180
gtgaagaaaa ttggcaaaagc tgtggaagac tccaagccct actgggagggc acggaggggtg 240
gcgaggcagg ctcagctgga agctcagaaa gccacgcagg acttccagag ggccacagag 300
gtgctccgtg ccgccaagga gaccatctcc ctggccgagc agcggctgct ggaggatgac 360
aagcggcgagt tcgactccgc ctggcaggag atgctgaatc acgccactca gagggtcatg 420
gaggcggagc agaccaagac caggagcgag ctgggtgcata aggagacggc agccaggtac 480
aatgccgcca tgggccgcat gcgacagctg gagaagaaac tcaagagagc catcaacaag 540
tccaagcctt attttgaact caaggcaaag tactatgtgc agctcgagca actgaaaaag 600
actgtggatg acctgcaggc caaactgacc ctggcaaaaag gcgagtacaa gatggccctg 660
aagaacctgg agatgatctc agatgagatc cacgagcggc ggcgctccag tgccatgggg 720
cctcggggat gcggtgtttg tgtgaggggc agcagcacat ctgtggagga tctgccaggg 780
agcaaacctg agcctgatgc catttctgtg gcctcggagg cctttgaaga tgacagctgt 840
agcaactttg tgtctgaaga tgactcggaa acccagtcctg tgtccagctt tagttcagga 900
ccaacaagcc cgtctgagat gcctgaccag ttccctgcgg ttgtgagggc tggcagcctg 960
gatctgcccc gccctgtgtc cctgtcagag tttgggatga tgttcccagt gttgggccct 1020
cgaagtgaat gcagcggggc ctccctccct gaatgtgaag tagaacgagg agacagggca 1080
gaaggggcag agaataaaac aagtgaacaa gccacaaca accggggcct cagcagtagc 1140
agtggcagtg gtggcagcag taagagccaa agcagcacct cccctgaggg ccaggccttg 1200
gagaaccgga tgaagcagct ctccctacag tgctcaaagg gaagagatgg aattattgct 1260
gacataaaaa tgggtcagat tggctgattc atcctggggc ctggccgatg tgcatatcaa 1320
catttatata tggaaactgga gaacattgtg ccaataatca tttaatatat gccaaatctt 1380
acacgtctac tctaaactgc tctaatagaag ttccagtgc cttgagggct aaagattggt 1440
cttctgggta agagctcttg ggctgggtttt tcagagcaga gttcttgttg tgggtagact 1500
gtgactaggt tcacagcctt tgtggaacat tccgtataac ggcattgttg aagcaataac 1560
tagttcctat gaaagaacca gagctgggaa gatggctggg aagccaggcc aaagtggggg 1620
caacagcttg cttctctttc tcttctcacc ctgagtttgt atgggaaaat ggagatgtcc 1680
tctccacttt atcccacgat atctaaatga aaaagaaaaga aaaccacac acaaagcaaa 1740
aactcaagta ttaagagcac atatttttga ccagtgagg gcttaaaaaa aaaaaaatcc 1800
aagaacacaa ttcattttca ccacctctgg tgttcagagg gggcttttaa aaaagcgtgt 1860
atgctgggat acccattaaa accattttct agaaggctac catgagctgc actttttggg 1920
gtgggaaagg tgaatgccag tggggatgag gggggatgag ggtaggaggg acttatgaa 1980
ggggatttatt ggtgtgggg gagaaaggtc tacagcataa gccttatcct gccagccaag 2040
gggatttatt ctaagagaag tgcatgtgaa gaatgggttc cactgttatt agattgacaa 2100
gatgttaatt tctctgtagg ttgtaacttt aaaaataaat gaaattattt aagggttatg 2160
ctgcactagt attccttaga ggaaacagtt ctttaaagtt aggaaaggga gtaggcaggc 2220
atgtgttggc aaaggctgtt aatagtagtt aagtgttaag actgcttttc tttaacgttt 2280
tcatggtaat gcatatttag agcactgtat tttgtcttg ttaagaaaat ttagcatttc 2340
taaaagaaaa aagcaaccct ctttcaaact gttaattctg tcacagcctg tatattttag 2400

```

tcattttgttaa atctcttcat acaatagtga cttctttttt gactgataca gtatcttaat 2460  
 tacaagggtta ttttgtactt gtcttaatac actaagtgtg ataaaaacgg cttgagaaaa 2520  
 gttaaaaaaa aaaa 2534

<210> 47

<211> 3786

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 2299715

<400> 47

ccgtcctcga ggcgaggaga gtaccggggc ggcccggtcg ccgcgcgagg agcgcggtcg 60  
 gcggcctggt ctgcggtga gatacacaga gcgacagaga catttattgt tatttgtttt 120  
 ttggtggcaa aaagggaaaa tggcgaacga ctcccctgca aaaagtctgg tgacatcga 180  
 cctctcctcc ctgcgggatc ctgctgggat ttttgagctg gtggaagtgg ttggaaatgg 240  
 cacctatgga caagtctata agggctcgaca tggtaaaaacg ggtcagttgg cagccatcaa 300  
 agttatggat gtcactgagg atgaagagga agaaatcaa ctggagataa atatgctaaa 360  
 gaaatactct catcacagaa acattgcaac atattatggt gctttcatca aaaagagccc 420  
 tccaggacat gatgaccaac tctggcttgt tatggagtgc tgtggggctg ggtccattac 480  
 agaccttgtg aagaacacca aagggaacac actcaaagaa gactggatcg cttacatctc 540  
 cagagaaatc ctgaggggac tggcacatct tcacattcat catgtgattc accgggatat 600  
 caagggccag aatgtgttgc tgactgagaa tgcaggggtg aaacttgttg actttgggtg 660  
 gagtgtctcag ctggacagga ctgtggggcg gagaaatagc ttcataaggca ctccctactg 720  
 gatggctcct gaggtcatcg cctgtgatga gaaccagat gccacctatg attacagaag 780  
 tgatcttttg tcttgtggca ttacagccat tgagatggca gaaggtgctc cccctctctg 840  
 tgacatgcac ccaatgagag cactgtttct cattcccaga aaccctcctc cccggctgaa 900  
 gtcaaaaaaa tggtcgaaga agtttttttag ttttatagaa ggggtgcttg tgaagaatta 960  
 catgcagcgg cctctacag agcagctttt gaaacatcct tttataaggg atcagccaaa 1020  
 tgaaggcaa gttagaatcc agcttaagga tcatatagat cgtaccagga agaagagagg 1080  
 cgagaagat gaaactgagt atgagtacag tgggagttag gaagaagagg aggaagtgcc 1140  
 tgaacaggaa ggagagccaa gttccattgt gaacgtgcct ggtgagtcta ctcttcgccc 1200  
 agatttcctg agactgcagc aggagaacaa ggaacgttcc gaggtctctc ggagacaaca 1260  
 gttactacag gagcaacagc tccggggagca ggaagaatat aaaaggcaac tgctggcaga 1320  
 gagacagaag cggattgagc agcagaaaaga acagaggcga cggctagaag agcaacaaag 1380  
 gagagagcgg gaagctagaa ggcagcagga acgtgaacag cgaaggagag aacaagaaga 1440  
 aaagaggcgt ctagaggagt tggagagaag gcgcaaagaa gaagaggaga ggagacgggc 1500  
 agaagaagaa aagaggagag ttgaaagaga acaggagtat atcaggcgac agctagaaga 1560  
 ggagcagcgg cacttggaag tctttcagca gcagctgtc caggagcagg ccatgttact 1620  
 gcatgaccat aggaggccgc acccgcagca ctgcagcag ccgccaccac cgcagcagga 1680  
 aaggagcaag ccaagcttcc atgctcccga gcccaaagcc cactacgagc ctgctgaccg 1740  
 agcgcgagag gttctgtga gaacaacatc tcgctcccct gttctgtccc gtcgagattc 1800  
 cccactgcag ggcaagtggc agcagaatag ccaggcagga cagagaaact ccaccagtat 1860  
 tgagcccagg cttctgtggg agagagtgga gaagctggtg cccagacctg gcagtggcag 1920  
 ctctcaggg tccagcaact caggatccca gcccggtct caccctgggt ctcagagtgg 1980  
 ctccggggaa cgcttcagag tgagatcatc atccaagtct gaaggctctc catctcagcg 2040  
 cctggaaaat gcagtgaaaa aacctgaaga taaaaaggaa gttttcagac ccctcaagcc 2100  
 tgctgactctg accgactgg ccaaagagct tcgagcagtg gaagatgtac ggccacctca 2160  
 caaagtaacg gactactcct catccagtga ggagtcgggg acgacggatg aggaggacga 2220  
 cgatgtggag caggaagggg ctgacgagtc cacctcagga ccagaggaca ccagagcagc 2280  
 gtcactctctg aatttgagca atggtgaaac ggaatctgtg aaaaccatga ttgtccatga 2340  
 tgatgtagaa agtgagccgg ccatgacccc atccaaggag ggcactctaa tcgtccgcca 2400  
 gactcagtc gctagtagca cactccagaa acacaaatct tcctcctcct ttacaccttt 2460

```

tatagacccc agattactac agattttctcc atctagcggg acaacagtga catctgtggt 2520
gggattttcc tgtgatggga tgagaccaga agccataagg caagatccta cccggaaagg 2580
ctcagtggttc aatgtgaatc ctaccaacac taggccacag agtgacaccc cggagattcg 2640
taaatacaag aagaggttta actctgagat tctgtgtgct gccttatggg gagtgaattt 2700
gctagtggtt acagagagtg gcctgatgct gctggacaga agtggccaag ggaaggtcta 2760
tcctcttata aaccgaagac gattttcaaca aatggacgta cttgagggct tgaatgtctt 2820
ggtgacaata tctggcaaaa aggataagtt acgtgtctac tatttgcctt ggtaagaaa 2880
taaaaatactt cacaatgatc cagaagttga gaagaagcag ggatggacaa ccgtagggga 2940
tttggaagga tgtgtacatt ataaagttgt aaaatatgaa agaatacaat ttctggtgat 3000
tgctttgaag agttctgtgg aagtctatgc gtgggcacca aagccatata acaaatttat 3060
ggcctttaag tcatttggag aattggtaca tggatcctgt gctggattcc atgctgttga 3120
gtgtgattca ggatcagttc atgacattta tctaccaaca catatccagt gtagcatcaa 3180
accccatgca atcatcatcc tccccaatat agatggaatg gagcttctgg tgtgctatga 3240
agatgagggg gtttatgtaa acacatatgg aaggatcacc aaggatgtag ttctacagt 3300
gggagagatg cctacatcag tagcatatat tcatccaat cagacaatgg gctggggaga 3360
gaaggccata gagatccgat ctgtggaaac tggctcattg gatggtgtgt tcatgcacaa 3420
aagggctcaa agactaaaat tcttgtgtga acgcaatgac aaggtgttct ttgcctctgt 3480
tcggtctggt ggagcagtc aggtttatct catgacctta ggcaggactt ctcttctgag 3540
ctggtagaag cagtgtgatc cagggattac tggcctccag agtcttcaag atcctgagaa 3600
cttggaattc cttgtaactg gagctcggag ctgcaccgag ggcaaccagg acagctgtgt 3660
gtgcagacct catgtgttgg gttctctccc ctcttctctg ttctctttat ataccagttt 3720
atccccattc tttttttttt tcttactcca aaataaatca aggttgcaat gcagctggtg 3780
ctgtta

```

&lt;210&gt; 48

&lt;211&gt; 1182

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 209854

&lt;400&gt; 48

```

gttggtgaag tcaagcgaag gcgactagag ctccaggagg gccagttctg tgggctctag 60
tcggccatat taataaagag aaaggggaagg ctgaccgtcc ttgcgctccg cccccacata 120
cacacccctt ctccccactc cgctctcacg actaagctct cagcatgaag gcacgcctgc 180
ctcgattgtc cagcctctgc cagaagaaag cttagcagcc agcgcctcag tagagacct 240
agggcgctga atgagtggga aagggaaatg ccgaccaatt gcgctgcggc gggctgtgcc 300
actacctaca acaagcacat taacatcagc ttccacaggg ttcttttggg tcctaaaaga 360
agaaaagaat gggttcgctt ggttaggcgc aaaaattttg tgccaggaaa acacactttt 420
ctttgttcaa agcactttga agcctcctgt tttgacctaa caggacaaac tcgacgactt 480
aaaatggatg ctgttccaac catttttgat tttgtacct atataaagtc tatgaaactc 540
aagtcaagga atcttttgaa gaaaaacaac agttgttctc cagctggacc atctaattta 600
aaatcaaaca ttagtagtca gcaagtacta cttgaacaca gctatgcctt taggaatcct 660
atggaggcaa aaaagaggat cattaaactg gaaaaagaaa tagcaagctt aagaagaaa 720
atgaaaactt gcctacaaaa ggaacgcaga gcaactcgaa gatggatcaa agccacgtgt 780
ttggtaaaga atttagaagc aaatagtgtg ttacctaaag gtacatcaga acacatgtta 840
ccaaactgct taagcagtc tcttttgaa gattttaaga tcttgaaca agatcaacaa 900
gataaaacac tgctaagtct aaatctaaaa cagaccaaga gtaccttcat ttaatttag 960
cttgcacaga gcttgatgcc tctcttcat tcttttcaga agtaaagata attatggcac 1020
ttatgccaaa attcattatt taataaagtt ttacttgaag taacattact gaatttgtga 1080
agacttgatt acaaaagaat aaaaaacttc atatggaaat tttatttgaa aatgagtggg 1140
agtgccttac attagaatta cggaacttca aaactatgat aa 1182

```

<210> 49  
 <211> 1676  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone Number: 1384286

<400> 49  
 tcgccgagcc cgtccgccgc cgccatggcc accacgggtga cctgcacccg cttcacccgac 60  
 gagtaccagc tctacgagga tattggcaag ggggctttct ctgtgggtccg acgctgtgtc 120  
 aagctctgca ccggccatga gtatgcagcc aagatcatca acaccaagaa gctgtcagcc 180  
 agagatcacc agaagctgga gagagaggct cggatctgcc gccttctgaa gcattccaac 240  
 atcgtgcgtc tccacgacag catctccgag gagggcttcc actacctggt cttcgatctg 300  
 gtcactgggtg gggagctctt tgaagacatt gtggcgagag agtactacag cgaggtgat 360  
 gccagtcact gtatccagca gatcctggag gccgttctcc attgtcacca aatgggggtc 420  
 gtccacagag acctcaagcc ggagaacctg cttctggcca gcaagtcaa aggggtgca 480  
 gtgaagctgg cagacttcgg cctagctatc gaggtgcagg gggaccagca ggcatgggtt 540  
 ggtttcgctg gcacaccagg ctacctgtcc cctgaggtcc ttcgcaaaga ggcgtacggc 600  
 aagcccggtg acatctgggc atgtgggggtg atcctgtaca tcctgctcgt gggctacca 660  
 cccttctggg acgaggacca gcacaagctg taccagcaga tcaaggctgg tgcctatgac 720  
 ttcccgctcc ctgagtggga caccgtcact cctgaagcca aaaacctcat caaccagatg 780  
 ctgaccatca accctgccaa gcgcatcaca gcccatgagg ccctgaagca cccgtgggtc 840  
 tgccaacgct ccacggtagc atccatgatg cacagacagg agactgtgga gtgtctgaaa 900  
 aagttcaatg ccaggagaaa gctcaaggga gccatcctca ccaccatgct ggccacacgg 960  
 aatttctcag cagccaagag tttactcaac aagaaagcag atggagtcaa gcccatacag 1020  
 aatagcacca aaaacagtgc agccgccacc agcccccagg ggacgcttcc tcctgccgcc 1080  
 ctggagtctt ctgacagtgc caataccacc atagaggatg aagacgctaa agcccgaag 1140  
 caggagatca ttaagaccac ggagcagctc atcgaggccg tcaacaacgg tgactttgag 1200  
 gcctacgcga aaatctgtga cccagggctg acctcgtttg agcctgaagc actgggcaac 1260  
 ctggttgaag ggtggactt ccacagattc tacttcgaga acctgctggc caagaacagc 1320  
 aagccgatcc acacgaccat cctgaaccca cagtgacag tcattggaga ggtgcccgc 1380  
 tgcacgctt acatccggct caccgagtac attgacgggc agggccggcc ccgcaccagc 1440  
 cagtctgagg agaccgcgt gtggcaccgc gcgcagcgca agtggcagaa cgtgcacttc 1500  
 cactgctcgg gcgcgcctgt ggccccgctg cagtgaagag ctgcgccctg gtttcgccgg 1560  
 acagagtgg tgttggagc ccgactgccc tcgggcacac ggcctgcctg tcgcatgttt 1620  
 gtgtctgcct cgttcctcc cctggtgcct gtgtctgcag aaaaacaagc ccgact 1676

<210> 50  
 <211> 1597  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone Number: 1512656

<400> 50  
 tcggccttcg gaaagacccc cgggccgggg cacggagaga gccgagcgcc gcagccgtga 60  
 gccgaataga gccggagaga cccgagtatg accggagaag cccaggccgg ccggaagagg 120  
 agccgagcgc ggccggaagg aaccgagccc gtccgaaggg agcggacgca gcctggcctg 180  
 gggccccggtc gagccccgc catggcgccc gaggcgacag ctgtggccgg aagcggggct 240  
 gttggcggtc gcctggccaa agacggcttg cagcagtcta agtggccgga cactacccca 300  
 aaacggcggc gcgcctcgtc gctgtcgcgt gacgccgagc gccgagccta ccaatggtgc 360

```

cgggagtact tgggcggggc ctggcgccga gtgcagcccg aggagctgag ggtttaccce 420
gtgagcggag gcctcagcaa cctgctcttc cgctgctcgc tcccggaaca cctgcccagc 480
gttggcgagg agccccggga ggtgcttctg cggctgtacg gagccatctt gcagggcgctg 540
gactccctgg tgctagaaaag cgtgatgttc gccatacttg cggagcggtc gctggggccc 600
cagctgtacg gagtcttccc agagggccgg ctggaacagt acatcccaag tcggccattg 660
aaaactcaag agcttcgaga gccagtgttg tcagcagcca ttgccacgaa gatggcgcaa 720
tttcatggca tggagatgcc tttcaccaag gagccccact ggctgttttg gaccatggag 780
cggtagctaa aacagatcca ggacctgccc ccaactggcc tccctgagat gaacctgctg 840
gagatgtaca gcctgaagga tgagatgggc aacctcagga agttactaga gtctaccca 900
tcgccagtcg tcttctgcca caatgacatc caggaaggga acatcttgct gctctcagag 960
ccagaaaatg ctgacagcct catgctggtg gacttcgagt acagcagtta taactatagg 1020
ggctttgaca ttgggaacca tttttgtgag tgggtttatg attatactca cgaggaatgg 1080
cctttctaca aagcaaggcc cacagactac cccactcaag aacagcagtt gcattttatt 1140
cgtcattacc tggcagaggc aaagaaagggt gagaccctct cccaagagga gcagagaaaa 1200
ctggaagaag atttctggtt agaagtcagt cggtatgctc tggcatccca tttcttctgg 1260
ggtctgtggt ccatcctcca ggcattccatg tccaccatag aatttggtta cttggactat 1320
gccagtcctc ggttccagtt ctacttccag cagaaggggc agctgaccag tgtccactcc 1380
tcactctgac tccacctcc cactccttgg atttctcctg gagcctccag ggcaggacct 1440
tggagggagg aacaacgagc agaaggccct ggcgactggg ctgagccccc aagtgaact 1500
gaggttcagg agaccggcct gttcctgagt ttgagtaggt ccccatggct ggcaggccag 1560
agccccgtgc tgtgtatgta acacaataaa caagctg 1597

```

&lt;210&gt; 51

&lt;211&gt; 2145

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2098635

&lt;400&gt; 51

```

cccacgcgtc cggacagctt gaccagttt gctttccaat caaagggcag ttattttgaa 60
tgtctctttg tggcgcaaga gccaacgcaa aaatgatggc ggcttacaat ggcggtacat 120
ctgcagcagc agcaggtcac caccaccacc atcaccacca ccttccacac ctccctctc 180
ctcacctgct tcaccaccac caccctcaac accatcttca tccggggctg gctgccgctg 240
tacacctgtg acagcagcac acctcttcgg cagctgcggc agccgcagca gcggctgcag 300
ctgcagccat gttaaaccct gggcaacaac agccatattt cccatcaccc gcaccggggc 360
aggctcctgg accagctgca gcagccccag ctgaggtaca ggctgccgca gctgctacag 420
ttaaggcgca ccatcatcag cactcgcac atccacagca gcagctggat attgagccgg 480
atagacctat tggatatgga gcctttggtg ttgtctggtc agtaacagat ccaagagatg 540
gaaagagagt agcgctcaaa aagatgccca acgtcttcca gaatctggtc tcttgcaaaa 600
gggtcttccg ggaattgaag atgttgtgtt tttttaagca tgataatgta ctctctgccc 660
ttgacatact ccaacctcca cacattgact attttgaaga aatatatgtt gtcacagaat 720
tgatgcagag tgacctacat aaaattatcg tctctcctca accactcagc tcagatcatg 780
tcaaagtttt tctttatcag attttgcgag gtttgaata tctccattca gctggcattt 840
tacatcgaga cattaagcca gggaatctcc ttgtgaacag caactgtgtt cttaaagattt 900
gtgatttttg attggccaga gtggaagagt tagatgaatc ccgtcatatg actcaggaag 960
ttgttactca gtattatcgg gctccagaaa tctgatggg cagccgtcat tacagcaatg 1020
ctattgacat ctggtctgtg ggatgtatct ttgcagaact actaggacga agaattatgt 1080
ttcaggcaca gagtcccatt cagcagttgg atttgatcac ggatctgttg ggcacaccat 1140
cactggaagc aatgaggaca gcttgtgaag gcgctaaggc acatatactc aggggtcctc 1200
ataaacagcc atctcttctt gtactctata ccctgtctag ccaggtetaca catgaagctg 1260
ttcatctcct ttgcaggatg ttggtctttg atccatccaa aagaatatcc gctaaggatg 1320
ccttagccca cccctaccta gatgaagggc gactacgata tcacacatgt atgtgtaaat 1380

```

```

gttgcttttc cacctccact ggaagagttt ataccagtga ctttgagcct gtcaccaatc 1440
ccaaatttga tgacactttc gagaagaacc tcagttctgt ccgacaggtt aaagaaatta 1500
ttcatcagtt cattttggaa cagcagaaag gaaacagagt gcctctctgc atcaaccctc 1560
agtctgctgc ttttaagagc tttattagtt ccactgttgc tcagccatct gagatgcccc 1620
catctcctct ggtgtgggag tgatgggtga agataatgta ctactgaaga tgtaatgtag 1680
ctttccactg gagtctggga tttgcaattc tggagggtta tcatgcttgt actgtaattt 1740
tactaatgaa gttttaaat aacaaccact acttgtatga tatgaataat atttagaaat 1800
gttactagac ttttaactct gtaaagtggg tgtgctttta gaagaaaaat attttaccce 1860
gagttgcaca tgttttatga atttagtgca gctgttatgg ctcacctcag aacaaaagag 1920
aattgaacca aatttgggag tttgggggtt tatgttttgt ttttcttttc taaaatgaag 1980
tgagattgtt cacacacaca cacacacaca cacacacaca cacaacaca aaggacagtc 2040
atacattttg atatttgagc cattcctaaa gatgtggggg tttctaaaac taaagaatct 2100
aggaaccttg cctgcgacca atcatggagc cacgtgagct gatcg 2145

```

<210> 52

<211> 1454

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 2446646

<400> 52

```

gggttcgaat tgcaacggca gctgccgggc gtatgtgttg gtgctagagg cagctgcagg 60
gtctcgctgg gggccgctcg ggaccaatth tgaagaggta cttggccacg acttattttc 120
acctccgacc tttccttcca ggcggtgaga ctctggactg agagtggctt tcacaatgga 180
agggatcagt aatttcaaga caccaagcaa attatcagaa aaaaagaaat ctgtattatg 240
ttcaactcca actataaata tcccggcctc tccgtttatg cagaagcttg gctttggtac 300
tggggtaaat gtgtacctaa tgaaaagatc tccaagaggt ttgtctcatt ctcttgggc 360
tgtaaaaaag attaatccta tatgtaatga tcattatcga agtgtgtatc aaaagagact 420
aatggatgaa gctaagattt tgaaaagcct tcatcatcca aacattgttg gttatcgtgc 480
ttttactgaa gccaatgatg gcagtctgtg tcttgctatg gaatatggag gtgaaaagtc 540
tctaaatgac ttaatagaag aacgatataa agccagccaa gatccttttc cagcagccat 600
aattttaaaa gttgctttga atatggcaag aggggttaaag tatctgcacc aagaaaagaa 660
actgcttcat ggagacataa agtcttcaaa tgttgtaatt aaaggcgatt ttgaaacaat 720
taaaatctgt gatgtaggag tctctctacc actggatgaa aatatgactg tgactgacct 780
tgaggcttgt tacattggca cagagccatg gaaacccaaa gaagctgtgg aggagaatgg 840
tgttattact gacaaggcag acatatttgc ctttggcctt actttgtggg aaatgatgac 900
tttatcgatt ccacacatta atctttcaaa tgatgatgat gatgaagata aaacttttga 960
tgaaagtgat tttgatgatg aagcatacta tgcagcggtg ggaactaggc cacctattaa 1020
tatggaagaa ctggatgaat cataccagaa agtaattgaa ctcttctctg tatgactaa 1080
tgaagaccct aaagatcgtc cttctgctgc acacattgtt gaagctctgg aaacagatgt 1140
ctagtatca tctcagctga agtgtggctt gcgtaaataa ctgtttattc caaaatattt 1200
acatagttac tatcagtagt tattagactc taaaattggc atatttgagg accatagttt 1260
cttggttaaca tatggataac tatttctaata atgaaatatg cttatattgg ctataagcac 1320
ttggaattgt actgggtttt ctgtaaagtt ttagaaacta gctacataag tactttgata 1380
ctgctcatgc tgacttaaaa cactagcagt aaaacgctgt aaactgtaac attaaattga 1440
atgaccatta cttt 1454

```

<210> 53

<211> 3225

<212> DNA

<213> Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 2764911

&lt;400&gt; 53

```

tggagcaggg ggcggtttgg ttgcgcggta ctagcgggtgc ccgccgaatg gggaggaggc 60
gaggagcgag ccgtgcggcc agagcgggaa agagactcgt ctttgcgtcc gagttctgga 120
gccgcgcac ccgactcctt ggggcgcggy cagcggctgc gaggggacgg gcgtccgctg 180
tctcctgggt tccccctgta gcgaccgcg ggatcgga aaaggagaa gatggaggag 240
gagggtggca gcagcggcgg cgccgcgggg accagcgcgg acggcggcga cggaggagag 300
cagctcctca ctgtcaagca cgagctgcgg actgctaatt tgacaggaca tgctgagaag 360
gtgggaatag aaaattttga gtcctgaag gtcctaggaa ctggagctta tggaaaagta 420
tttctagttc gtaaaataag tggccatgat actggaaagc tgtatgccat gaaagttttg 480
aaaaaggcaa caatcgttca aaaggccaaa accacagagc atacaaggac agaacgacaa 540
gtcctggaac acattaggca gtgcgcattt ttggtaacat tacattatgc tttccagaca 600
gaaaccaaac ttcactctcat tttagattat ataaatggtg gtgaactttt taccatctt 660
tctcaaagag agcgtttcac agagcatgag gtgcagattt atgttggaga gattgtgctt 720
gccctcgaac atctccacaa gttggggatt atatatcgtg atattaagct tgagaatatt 780
ctacttgatt ctaatggcca tgtggtgctg acagattttg gtctgagtaa ggagtttggtg 840
gctgatgaaa ctgaaagagc atattccttt tgtggaacta ttgaatacat ggcaccagat 900
attgtcagag ggggagattc aggcacatgac aaggcagttg actggtggag tttgggtggt 960
ctaattgtatg aattactaac tggagcatct cctttcactg ttgatggaga aaaaaattcc 1020
caagctgaga tatctaggag aatattaaaa agtgagcctc catatcccca agaaatgagt 1080
gcttttagcga aagaccta atcagcgtctt ttgatgaaag atcccaagaa gagattggga 1140
tgtggtccac gtgatgcaga tgaaatcaaa gaacatctct tctttcagaa aataaattgg 1200
gatgatattag ccgcaaaaaa agtgccctgca ccatttaagc cagtcattcg agatgaatta 1260
gatgtgagta actttgcaga agagttcaca gaaatggatc ccacttattc tccgcagacc 1320
ctgccccaga gttctgagaa gctgtttcag ggctattcct ttgttgcctc ttccatccta 1380
ttcaagcgta atgcagctgt catagaccct cttcagtttc acatgggagt tgaacgtcct 1440
ggagtgcaca atgttgccag gagtgcaatg atgaaggact ctccattcta tcaacactat 1500
gacctagatt tgaaggacaa acccctggga gaaggtagtt tttcaatttg tcaaaagtgt 1560
gtgcataaaa aaagtaacca agcttttgca gtcaaaataa tcagcaaaag gatggaagcc 1620
aataactcaa aggaaataac agctctggaa ctctgtgaag gacaccccaa tattgtgaag 1680
ttgcatgaag tttttcatga tcagcttcac acgtttctag tgatggaact tctgaatgga 1740
ggagaactgt ttgagcgcag taagaaaaag aagcacttca gtgagacgga agccagctac 1800
atcatgagga agcttgtttc agctgtaagc cacatgcatg atgttggagt ggtgcacagg 1860
gatctgaaac ctgagaattt attgttcacc gatgaaaatg acaatttgga aattaaaata 1920
attgattttg gatttgcacg gctaaagcca ccggataatc agccctgaa gactccatgc 1980
ttcacccctc attatgccg cccagagctc ttgaatcaga acggctacga tgagtctgt 2040
gacctgtgga gcttgggctg cattttgtac acaatgttgt caggacaggt tcccttccaa 2100
tctcatgacc gaagtttgac gtgtaccagc gcggtggaaa tcatgaagaa aattaaaaag 2160
ggagattttc ctttgaagg agaagcctgg aagaatgtat cccaagaggc taaagatttg 2220
atccaaggac ttctcacagt agatccaaac aaaaggctta aaatgtctgg cttgaggtac 2280
aatgaatggc tacaagatgg aagtcagctg tcctccaatc ctctgatgac tccggatatt 2340
ctaggatctt ccggagctgc cgtgcatacc tgtgtgaaag caaccttcca cgcctttaac 2400
aaatacaaga gagaggggtt ttgccttcag aatgttgata agggcccttt ggctaagaga 2460
agaaaaatga aaaagactag caccagtacc gagacacgca gcagttccag tgagagttcc 2520
cattcttctt cctctcattc tcacggtaaa actacacca ccaagacact gcagcccagc 2580
aatcctgccg acagcaataa cccggagacc ctcttcaggt tctcggactc agtagcttag 2640
gcatggtagg agtgtatcag tgatccattg cacctttatt ccctcagcat atgcctgagg 2700
cgatctttta tgcttttaaa aatgtttccc gttggtctca ttggaatctg cctcctaatt 2760
atttttttca ggaaaacctg tttggttatc ctcatcaca agcactggac agagaatgtt 2820
actgtgaata gagcacatat tactcttttt agcaacctag catgatgcca acaagactat 2880
tcttgaaaga gcaaaggttc ctgtaaat tt aattagggct agatttgagc tgcttgtaag 2940
tcacagggtt tccagatgtc tgccaacaag aaatgactca tactgtgatg ataccttttg 3000
ctttgccttg tggacaatgt gggtttttga aatttgcacc cttcaaacaa tgatttatca 3060

```

gagaaagggg tctgttttca aaaaagattc tgtaatgaat tttatgtgtg gcatatactt 3120  
 atttcttgag agaagatttt aacttattgt ttttatttta tggttacata tgatgataac 3180  
 ctgctattat taaacttttt ctaaaaagtg aaaaaaaaaa aaaaa 3225

<210> 54

<211> 2110

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 3013946

<400> 54

tcgccgagcc cgtccgccgc cgccatggcc accacgggtga cctgcaccgc cttcacccgac 60  
 gagtaccagc tctacgagga tattggcaag ggggctttct ctgtgggtccg acgctgtgtc 120  
 aagctctgca cgggccatga gtatgcagcc aagatcatca acaccaagaa gctgtcagcc 180  
 agagatcacc agaagctgga gagagaggct cggatctgcc gccttctgaa gcattccaac 240  
 atcgtgcgtc tccacgacag catctccgag gagggcttcc actacctggt cttcgatctg 300  
 gtcactggtg gggagctctt tgaagacatt gtggcgagag agtactacag cgaggctgat 360  
 gccagtcaat gtatccagca gatcctggag gccgttctcc attgtcacca aatgggggtc 420  
 gtccacagag acctcaagcc ggagaacctg cttctggcca gcaagtgcaa aggggctgca 480  
 gtgaagctgg cagacttcgg cctagctatc gaggtgcagg gggaccagca ggcgtatggt 540  
 ggtttcgtcg gcacaccagg ctacctgtcc cctgaggtcc ttcgcaaaga ggcgtatggc 600  
 aagcctgtgg acatctgggc atgtgggggtg atcctgtaca tcctgctcgt gggctacca 660  
 cctctctggg acgaggacca gcacaagctg taccagcaga tcaaggctgg tgcctatgac 720  
 ttcccgctcc ctgagtggga caccgtcact cctgaagcca aaaacctcat caaccagatg 780  
 ctgaccatca accctgcca ggcgcacaca ccccatgagg cctgaagca cccgtgggtc 840  
 tgccaacgct ccacggtagc atccatgatg cacagacagg agactgtgga gtgtctgaaa 900  
 aagttcaatg ccaggagaaa gctcaaggga gccatcctca ccaccatgct ggccacacgg 960  
 aatttctcag ccaagagttt actcaacaag aaagcagatg gagtcaagcc ccagacgaat 1020  
 agcaccaaaa acagtgcagc cgccaccagc cccaaaggga cgcttcctcc tgccgccctg 1080  
 gagcctcaaa ccaccgtcat ccataaccca gtggacggga ttaaggagtc ttctgacagt 1140  
 gccaatacca ccatagagga tgaagacgct aaagcccccga ggggtcccga catcctgagc 1200  
 tcagtgagga ggggctcggg agccccagaa gccgaggggc ccctgcctcg cccatctccg 1260  
 gtcctcttg gcccctgcc agctccatcc cccagatct ctgacatcct gaactctgtg 1320  
 agaaggggtt caggaacccc agaagccgag gggccctct cagcggggcc cccgctctgc 1380  
 ctgtctccgg ctctcctagg cccctgtcc tcccgtccc ccaggatctc tgacatcctg 1440  
 aactctgtga ggaggggctc agggacccca gaagccaagg gcccctcgcc agtggggccc 1500  
 ccgcccctgc catctccgac tatccctggc cccctgccc ccccatccc gaagcaggag 1560  
 atcattaaga ccacggagca gtcacatgag gccgtcaaca acggtgactt tgaggcctac 1620  
 gcgaaaatct gtgaccagag gctgacctcg tttgagcctg aagcactggg caacctgggt 1680  
 gaagggatgg acttccacag attctacttc gagaacctgc tggccaagaa cagcaagcca 1740  
 atccacacga ccacctgaa cccacacgtg cacgtcattg gagaggatgc cgcctgcac 1800  
 gcttacatcc ggctcacgca gtacattgac gggcagggcc ggccccgcac cagccagtct 1860  
 gagagaccc gcgtgtggca ccgcccgcac ggcaagtggc agaattgtgca cttccactgc 1920  
 tcggggcgcc ctgtggcccc gctgcagtga agagctgcgc cctggtttcg ccggacagag 1980  
 ttggtgtttg gagcccact gccctcgggc acacggcctg cctgtcgcat gtttgtgtct 2040  
 gcctcgttcc ctcccctggg gcctgtgtct gcagaaaaac aagaccagat gtgatttgtt 2100  
 aaaaaaaaaa 2110

<210> 55

<211> 2140

<212> DNA



<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 067967

<400> 55

```

gtgcgctgag ctgcagtgtc tggtcgagag taccctgtgg agcgtcgcgc cgcgaggagca 60
gccgtcccgg cgtaggtggc gtggccgacc ggacccccaa ctggcgccctc tccccgcgcg 120
gggtcccagag ctaggagatg ggaggcacag ctctgtgggc tgggcggaag gatgcggggc 180
cgcttggggc cgggtctccg ccccgagcgc ggaggttggg ggatggtgtc tatgacacct 240
tcatgatgat agatgaaacc aaatgtcccc cctgttcaaa tgtactctgc aatccttctg 300
aaccaccttc acccagaaga ctaaatatga ccactgagca gtttacagga gatcatactc 360
agcacttttt ggatggaggt gagatgaagg tagaacagct gtttcaagaa tttggcaaca 420
gaaaatccaa tactattcag tcagatggca tcagtgactc tgaaaaatgc tctcctactg 480
tttctcaggg taaaagttca gattgcttga atacagtaaa atccaacagt tcatccaagg 540
caccctaaag ggtgcctctg actccagaac aagccctgaa gcaatataaa caccacctca 600
ctgcctatga gaaactggaa ataattaatt atccagaaat ttactttgta ggtccaaatg 660
ccaagaaaag acatggaggt attggtgggt ccaataatgg agggatgat gatgcagatg 720
gggcctatat tcatgtacct cgagaccatc tagcttatcg atatgaggtg ctgaaaatta 780
ttggcaaggg gagttttggg caggtggcca ggtctatga tcacaaactt cgacagtacg 840
tggccctaaa aatggtgcgc aatgagaagc gctttcatcg tcaagcagct gaggagatcc 900
ggatttttga gcatcttaag aaacaggata aaactggtag tatgaacgtt atccacatgc 960
tggaaagttt cacattccgg aacctgttt gcatggcctt tgaattgctg agcatagacc 1020
tttatgagct gattaaaaaa aataagtttc aggggttttag cgtccagttg gtacgcaagt 1080
ttgccagtc catcttgcaa tctttggatg cctccacaa aaataagatt attcactgcg 1140
atctgaagcc agaaaacatt ctctgaaac accacgggcg cagttcaacc aaggtcattg 1200
actttgggtc cagctgtttc gagtaccaga agctctacac atatatccag tctcggttct 1260
acagagctcc agaatcatc tttaggaagc gctacagcac accaattgac atatggagtt 1320
ttggctgcat ccttgcaaaa cttttaacag gacagcctct cttccctgga gaggatgaag 1380
gagaccagtt ggcctgcatg atggagcttc tagggatgcc accacaaaaa cttctggagc 1440
aatccaaacg tgccaaagta tttattaatt ccaagggcat accccgctac tgctctgtga 1500
ctaccagggc agatgggagg gttgtgcttg tgggggggtc ctcacgtagg ggtaaaaagc 1560
gggggtcccc aggcagcaaa gactggggga cagcactgaa aggggtgtgat gactacttgt 1620
ttatagagtt cttgaaaagg tgtcttcaat gggacccctc tgcccgttg accccagctc 1680
aagcattaaag acacccttg attagcaagt ctgtccccag acctctcacc accatagaca 1740
aggtgtcagg gaaacgggta gttaatcctg caagtgtctt ccagggttg ggttccaagc 1800
tgcctccagt tggtggaata gccataaagc ttaaagctaa cttaatgtca gaaaccaatg 1860
gtagtatacc cctatgcagt gtattgcaa aactgattag ctagtggaca gagatagcc 1920
cagagatgca tatgtgtata ttttatgat cttacaaacc tgcaaatgga aaaaatgcaa 1980
gccattgggt ggatgttttt gtttagagtag acttttttta aacaagacaa aacattttta 2040
tatgattata aaagaattct tcaagggcta attacctaac cagcttgat tggccatctg 2100
gaatatgcat taaatgactt tttataggtc aaaaaaaaaa 2140

```

<210> 56

<211> 1728

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 346275

<400> 56

```

gacagacaaa gcgccgccac gcgtccgcat gtcggatgtt tgtagcagtc agagagcaga 60

```

```

acatgagcat ctgccaggtc tgggtccccc accatcaggg atgggagtga gaaaggggag 120
ttcccctctg aagagccacc cctgcaggga gaaatctgtc tccaacagga gatctgggaa 180
gaccatagtg agaagtgtg tcgaagaggt ccgcacagcg ggccttttcc gaagtgggtt 240
tagcgaagag aaggcaactg gcaagctctt tgctgtgaag tgtatcccta agaaggcgct 300
gaaaggcaag gaaagcagca tagagaatga gatagccgtc ctgagaaaga ttaagcatga 360
aaatattgtt gccctggaag acatttatga aagcccaa atcacctgtact tggatcatga 420
gctgggtgtc ggtggagagc tgtttgaccg gatagtggag aaggggtttt atacagagaa 480
ggatgccagc actctgatcc gccaaagtct ggacgccgtg tactatctcc acagaatggg 540
catcgccac agagacctca agcccga aaa tctctgtac tacagtcaag atgaggagtc 600
caaaataatg atcagtgact ttggattgtc aaaaatggag ggcaaaggag atgtgatgtc 660
cactgcctgt ggaactccag gctatgtcgc tctgaagtc ctgcccaga aaccttacag 720
caaagccgtt gactgctgtt ccatcgaggt gattgcctac atcttgcctc gcggctacc 780
tcctttttat gatgaaaatg actccaagct ctttgagcag atcctcaagg cggaatatga 840
gtttgactct ccctactggg atgacatctc cgactctgca aaagacttca ttcggaacct 900
gatggagaag gacccgaata aaagatacac gtgtgagcag gcagctcggc acccatggat 960
cgctgggtgac acagccctca acaaaaacat ccacgagtc gtcagcgccc agatccggaa 1020
aaactttgcc aagagcaaat ggagacaagc atttaatgcc acggccgtcg tgagacatat 1080
gagaaaacta cacctcggca gcagcctgga cagttcaaat gcaagtgttt cgagcagcct 1140
cagtttgccc agccaaaaag actgtgcgta tgtagcaaaa ccagaatccc tcagctgaca 1200
ctgaagacga gcctgggggtg gagaggagg agccggcctc tgccgagcac ctctgtttg 1260
ccaggcgctt tctatactta atcccctgtc atgcgaccct aggactttt ttaacatgta 1320
atcactgggc cgggtgcagt ggctcacgcc tgtaatccca acactttggg aggtgaggg 1380
aggaggactg tttgagttca ggagttttaa gcacagcctg accaacaagg tgaaacccca 1440
tctctactaa aatataaaaa tttagccgggt gtgggtggcga gcacctgtaa tgtcagctac 1500
ttgggaggct gaggcaggag aatcacttga acccaggaag cggagggtgc aatgagctga 1560
gatcacacca ctgcactcca gcctgggtga cagattgaga ctccctctca aaaaaaaaag 1620
ggaaatcatt gaacactcgt ggaaccctag gtattgcata ttccatttac ggtttgggaa 1680
tccagggctc aagtcctcgc aggggtaccg agctcgagat cgtaatca 1728

```

&lt;210&gt; 57

&lt;211&gt; 1610

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 283746

&lt;400&gt; 57

```

gtcgcctctg aaggagaacc attttccatc tctttcatag ttttttcccc cagtcagcgt 60
ggtagcggta ttctccgagg cagtgcaggt aattgttttt gcctcttttag ccaagacttc 120
cgccctcgat caagatggtg gttggacggc cttcctaacc tttacggggc ctggcggtgc 180
tgacgcctga gctggtaggg gtggagcagg taggaaacag caaatgcaga agctgtctgcg 240
cggaagtcgg ccatggactg gaaagaagtt cttcgtcggc gcctagcgac gcccaacacc 300
tgtccaaaca ctgcctgctg aagatgaagt cttactacag aaattaagag aggaatcaag 360
agctgtcttt ctacaaaaga aaagcagaga actgttagat aatgaagaat tacagaactt 420
atggtttttg ctggacaaac accagacacc acctatgatt ggagaggaag cgatgatcaa 480
ttacgaaaac tttttgaagg ttggtgaaaa ggctggagca aagtgcagc aattttcac 540
agcaaaagtc tatgtctaat tccttcatac agattcatat ggaagaattt ccatcatgca 600
gttctttaat tatgtcatga gaaaagtttg gcttcatcaa acaagaatag gactcagttt 660
atatgatgtc gctgggcagg ggtaccttcg ggaatctgat ttagaaaact acatattgga 720
acttatccct acgttgccac aattagatgg tctggaaaaa tctttctact ctttttatgt 780
ttgtacagca gttaggaagt tcttctctct tttagatcct ttaagaacag gaaagataaa 840
aattcaagat atttttagcat gcagcttctt agatgattta ttggagctaa gggatgagga 900

```

```

actgtccaag gagagtcaag aaacaaattg gttttctgct ctttctgccc taagagttta 960
tggccagtac ttgaatcttg ataaagatca caatggcatg ctcaagtaaag aagaactctc 1020
acgctatgga acagctacca tgaccaatgt cttcttagac cgtgttttcc aggagtgtct 1080
cacttatgat ggagaaatgg actataagac ctacttgac tttgtccttg cattagaaaa 1140
cagaaaggaa cctgcagctc tacaatatat tttcaaactg cttgatattg agaacaaagg 1200
atacctgaat gtcttttcac ttaattattt ctttagggcc atacaggaac taatgaaaat 1260
ccatggacaa gatcctgttt catttcaaga tgtcaaggat gaaatctttg acatggtaaa 1320
accaaaggat cctttgaaaa tctctcttca ggatttaatc aacagtaatc aaggagacac 1380
agtaaccacc attctaactg atttgaatgg cttctggact tacgagaaca gagaggctct 1440
tgttgcaaat gacagtgaag actctgcaga ccttgatgat acatgatctc tgaaagacta 1500
gactgtctta tattatgaga tacttgaatg ctgcatgtaa agcctttaa gcaaaatcct 1560
cagaaatggt ctaaataaaa cacttgatat gcctagagaa aaaaaaaaaa 1610

```

<210> 58

<211> 1290

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 2696537

<400> 58

```

ccggctcccg ccgggaagtt ctaggccgcc gcacagaaag ccctgccctc cacgccgggt 60
ctctggagcg ccctgggttg cccggccggt ccctgccgct gacttggtga cactgcgagc 120
actcagtcct tcccgcgcgc ctccctcccg cccgccccgc cgctcctect ccctgtaaca 180
tgccatagtg cgctgcgac cacacggcgg gggcgctagc gttcgcttc agccaccatg 240
gggaatggga tgaacaagat cctgcccggc ctgtacatcg gcaacttcaa agatgccaga 300
gacgcggaac aattgagcaa gaacaagggt acacatatc tgtctgtcca tgatagtgcc 360
aggcctatgt tggagggagt taaatacctg tgcacccag cagcggattc accatctcaa 420
aacctgacaa gacatttcaa agaaagtatt aaattcattc acgagtgcg gctccgcggg 480
gagagctgcc ttgtacactg cctggccggg gtctccagga gcgtgacact ggtgatcgca 540
tacatcatga ccgtcactga ctttggttgg gaggatgcc tgcacaccgt gcgtgctggg 600
agatcctgtg ccaaccccaa cgtgggcttc cagagacagc tccaggagtt tgagaagcat 660
gaggtccatc agtatcggca gtggctgaag gaagaatatg gagagagccc tttgcaggat 720
gcagaagaag ccaaaaacat tctggccgct ccgggaattc tgaagttctg ggcctttctc 780
agaagactgt aatgtacctg aagtttctga aatattgcaa acccacagag tttaggctgg 840
tgctgccaaa aagaaaagca acatagagtt taagtatcca gtagtgattt gtaaaacttg 900
ttttcatttg aagctgaata tatacgtagt catgtttatg ttgagaacta aggatattct 960
ttagcaagag aaaatatttt ccccttatcc ccactgctgt ggaggtttct gtacctcgct 1020
tggatgctg taaggatccc gggagccttg ccgcactgcc ttgtgggtgg cttggcgctc 1080
gtgattgctt cctgtgaacg cctcccaagg acgagcccag tgtagtgtg tggcgtgaac 1140
tctgcccggt tggtctcaaa ttccccagct tgggaaatag ccctgggtgt gggttttatc 1200
tctggtttgt gttctccgtg gtggaattga ccgaaagctc tatgttttct ttaataaagg 1260
gcaacttagc caagtttaaa aaaaaaaaaa 1290

```

<210> 59

<211> 2281

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 551178

&lt;400&gt; 59

```

tgatgatcca gatgttaaag cacaagtggg agtgcgtgcc gctgcactac gtgcttccag 60
cctggatgca catgaagaga ccatcagtat agaaaagaga agtgatttgc aagatgaact 120
ggatataaat gagctaccaa attgtaaaat aaatcaagaa gattctgtgc ctttaatcag 180
cgatgctgtt gagaatattg actccactct tcatatatt cacagcgatt cagacttgag 240
caacaatagc agtttttagc ctgatgagga aaggagaact aaagtacaag atgttgtacc 300
tcaggcggtt ttagatcagt atttatctat gactgaccct tctcgtgcac agacgggtga 360
cactgaaatt gctaagcact gtgcatatag cctccctggg gtggccttga cactcggaag 420
acagaattgg cactgcctga gagagacgta tgagactctg gcctcagaca tgcagtggaa 480
agttcgacga actctagcat tctccatcca cgagcttgca gttattcttg gagatcaatt 540
gacagctgca gatctggttc caatttttaa tggattttta aaagacctcg atgaagtcag 600
gatagggtgt cttaaacact tgcattgatt tctgaagctt cttcatattg acaaaagaag 660
agaatatctt tatcaacttc aggagttttt ggtgacagat aatagtagaa attggcggtt 720
tcgagctgaa ctggctgaac agctgatttt acttctagag ttatatagtc ccagagatgt 780
ttatgactat ttacgtccca ttgctctgaa tctgtgtgca gacaaagttt cttctgttcg 840
ttggatttcc tacaagttgg tcagcgagat ggtgaagaag ctgcacgcgg caacaccacc 900
aacgttcgga gtggacctca tcaatgagct tgtggagaac tttggcagat gtcccaagtg 960
gtctggctcg caagcctttg tctttgtctg ccagactgtc attgaggatg actgccttcc 1020
catggaccag tttgctgtgc atctcatgcc gcactctgta accttagcaa atgacagggg 1080
tcctaacgtg cgagtgtgct ttgcaaagac attaagacaa actctactag aaaaagacta 1140
tttcttgccc tctgccagct gccaccagga ggctgtggag cagaccatca tggctcttca 1200
gatggaccgt gacagcgatg tcaagtattt tgcaagcatc caccctgcca gtacaaaat 1260
ctccgaagat gccatgagca cagcgtcctc aacctactag aaggcttgaa tctcgggtgc 1320
tttctgctt ccatgagagc cgagggttcag tgggcattcg ccacgcattg gacctgggat 1380
agctttcggg ggaggagaga ccttcctctc ctgcggactt cattgcaggt gcaagttgcc 1440
tacacccaat accagggatt tcaagagtca agagaaagta cagtaaaccac tattatctta 1500
tcttgacttt aaggggaaat aatttctcag aggattataa ttgtcaccga agccttaaatt 1560
ccttctgtct tctgactga atgaaacttg aattggcaga gcattttcct tatggaaggg 1620
atgagattcc cagagacctg cattgctttc tctgggtttt atttaacaat cgacaaatga 1680
aattcttaca gcctgaaggc agacgtgtgc ccagatgtga aagagacctt cagtatcagc 1740
cctaactctt ctctcccagg aaggacttgc tgggctctgt ggccagctgt ccagcccagc 1800
cctgtgtgtg aatcgtttgt gacgtgtgca aatgggaaaag gaggggtttt tacatctcct 1860
aaaggacctg atgccaacac aagtaggatt gacttaaact cttaagcgca gcatattgct 1920
gtacacattt acagaatggt tgcgtagtgt ctgtgtctga ttttttcatg ctggtcatga 1980
cctgaaggaa atttattaga cgtataatgt atgtctggtg tttttaactt gatcatgac 2040
agctctgagg tgcaacttct tcacatactg tacatactg tgaccactct tgggagtgct 2100
gcagtcttta atcatgctgt ttaaactgtt gtggcacaaag ttctcttctc caaataaaat 2160
ttattaataa gatctataga gagagatata tacacttttg attgttttct agatgtctac 2220
caataaatgc aatttgtgac ctgtattaat gatttaaaag gggaaactag attaaaatat 2280
a

```

&lt;210&gt; 60

&lt;211&gt; 632

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone Number: 619292

&lt;400&gt; 60

```

cggacgcgtg gggccagacc gcagctccag caccgaggac ttctgctacg tcttcacggg 60
ggagctggaa cgaggccctt ccgggctggg gatgggcctg atcgacggga tgcacacgca 120
cctgggcgcc ccgggctctt acatccagac cctgctcccg ggcagccccg cagcgccgca 180
cgggcgcctg tcgctggggg accgtatcct ggaggtgaat ggcagcagcc tctgggcct 240

```

```

tggtacctg agagctgtgg acctgatccg tcatggcggg aagaagatgc ggttcctggt 300
cgcgaaagtc gacgttggga aacagccaag aagatccatt tccgcacgcc ccctctctag 360
gggggctgcg aggacacccc cacaggcccg gcaccgggtc ccacctgggtg acactgggct 420
tcctcccgc ttcgtccctg ttttgtaact gaccaagttg ggtcccgggt ggggagcctc 480
accctgggga catgcctgtt gataacatgc atctcagttg aggttctatt tatatggcag 540
atgacgtgaa attgtgatgt ttgttacaga gcttttatgt ttaaagactt caatggagaa 600
gtacgggtca ataaactatt tttcccggtc tt                                     632

```

<210> 61

<211> 2347

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 2054049

<400> 61

```

cccagtttta tcatggattc atcctgaaag tcaagccaca atcactcgggt gtagccagcc 60
catggttgga gtgagtggaa agcgaagcaa agaagatgaa aaataccttc aagctatcat 120
ggattccaat gccagctctc acaaaatctt tatatttgat gcccggccaa gtgttaatgc 180
tgttgccaac aaggcaaagg gtggaggtta tgaagtgaa gatgcctatc aaaatgctga 240
actagttttc ctggatatcc acaatattca tgttatgaga gaatcattac gaaaacttaa 300
ggagattgtg tacccaaca ttgaggaaac ccactgggtg tctaacttgg aatctactca 360
ttggctagaa catattaagc ttattcttgc aggggctctt aggattgctg acaaggtaga 420
gtcagggaa acgtctgtgg tagtgcatg cagtgatggg tgggatcgca cagctcagct 480
cacttccctt gccatgctca tgttggatgg atactatcga accatccgag gatttgaagt 540
ccttgtggag aaagaatggc taagttttgg acatcgattt caactaagag ttggccatgg 600
agataagaac catgcagatg cagacagatc gcctgttttt cttcaattta ttgactgtgt 660
ctggcagatg acaagacagt ttcctaccgc atttgaattc aatgagtatt ttctcattac 720
cattttggac cacctataca gctgcttatt cggaacattc ctctgtaata gtgaacaaca 780
gagaggaaaa gagaatcttc ctaaaaggac tgtgtcactg tggctttaca taaacagcca 840
gctggaagac ttcactaatc ctctctatgg gagctattcc aatcatgtcc tttatccagt 900
agccagcatg cgccacctag agctctgggt gggatattac ataagggtgga atccacggat 960
gaaaccacag gaacctattc acaacagata caaagaactt cttgctaacc gagcagagct 1020
tcagaaaaaa gttagaggaac tacagagaga gatttctaac cgatcaacct catcctcaga 1080
gagagccagc tctcctgcac agtgtgtcac tctgtccaa actgttgtat aaaggactgt 1140
aagatcaggg gcatcattgc tatacactct tgattacact ggcagctcta tgagtagaaa 1200
gtcttcggaa tttagaacct atctatgaga gaaagttcag tcaactttatt tattttaaat 1260
ctctctagga tgagtttaga actgtagcag tgcagggtgc ttaagtgaag taactccata 1320
tgtaattaca tgattatgat actaatcttt taagtatcca aagaatatta aaatacttca 1380
atcctggatt cacagtggga acaagtttct attaaaaggc aaatgctgtt acaaattttt 1440
ggcatctggt aatattaaaa ccatttttaga aatacactct gtgctcactg tgcagaggaa 1500
catcagtttt caaaccaaca ctgaaattct gtggcatcac atatattggg ccttgatgtc 1560
atgacagatc aaaatcattt gatatccctt tctccattct aggtttttct ttttttcagt 1620
aactgattta ccttgatcac ttttcaactt ccatattctt catatagtaa aaggcaaagt 1680
gttgaagata ctacgggtgtg gttagtagttg aaaattattg ccgtcattat ttacatactt 1740
aagacatatt agcaagttga tccaaaatgg gaggccttat agatgtgctt gggggaaaaa 1800
gaaggggaga aagtagccat acaggagtgc aaagaattcc atgcccttca gattagccca 1860
attaccagaa acatcatgaa agatatttta aaaactaatt atttactaca gtgtatttca 1920
cttgtcttgt gtgtctgaac acacagaagc taattagcaa gtttttaaga agtattttaa 1980
aatcttacta ggattgacat tttttctgaa ttctgtataa atagcttata gtgagaagta 2040
ctgtgtctca attttacatt tttttccttt gcaaattctg taatttcact caacgattaa 2100
gtctaccaa gaacacactg catgtaaaag atgtattaca atctcaaagc cagtataaaga 2160
aatcttgctt cactgttcac ctgtctacaag taagagtttg gtgctggtag aaacatttga 2220

```

ctctgatgtc tattttattc tacataagag ccatatgtaa tgtactgtaa caaaggagct 2280  
 tcttgtcccc ttggtctttt aattaaaaga aattccaact gactttttaa ctttaaaaaa 2340  
 aaaaaaa 2347

<210> 62

<211> 1737

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte Clone Number: 2843910

<400> 62

ccggggctga gcgctcggct gcagcggcgc ggaggccgtc tccctggctc gccgcggctcc 60  
 ccgcccgtcc cgcgcgcggc tgccatggca ggagccggag ggttcggctg ccccgccgggc 120  
 ggcaacgact tccagtgggtg cttctcgag gtcaaggggg ccatcgacga ggacgtggcc 180  
 gaagcggaca tcatttccac cgttgagttt aattactctg gagatcttct tgcaacagga 240  
 gacaagggcg gcagagttgt tatttttcag cgtgaacaag agaataaaag ccgcctcat 300  
 tctaggggag aatataatgt ttacagcacc tttcaaagtc atgaaccgga gtttgactat 360  
 ttgaaaagtc tagaaattga ggaaaaaatt aataaaatta ggtggttacc acaacagaat 420  
 gctgctcatt ttctactgtc tacaatgat aaaactataa aattatggaa aataagtga 480  
 cgggataaaa gagcagaagg ttataacctg aaagacgaag atggaagact tcgagaccca 540  
 tttaggatca cggcgtacg ggtcccaata ttgaagccca tggatcttat ggtagaagcg 600  
 agtccacggc gaatttttgc aaatgctcac acatatcata taaattccat ttcagtaaat 660  
 agtgatcatg aaacatatct ttctgcagat gacctgagaa ttaatttatg gcacttagaa 720  
 atcacagata gaagctttaa catcgtggac atcaagcctg ctaacatgga ggagctgacc 780  
 gaagtcacga ctgcagccga gttccaccgc caccagtgc acgtgttcgt ctacagcagt 840  
 agcaaaagga ccatccgcct gtgtgacatg cgctcctcgg ccctgtgcga cagacactcc 900  
 aagttttttg aagagcctga agatcccagc agtaggtcct tcttctcaga aataatttca 960  
 tccatatccg atgtaaaatt cagtcatagt gggcgggtaca tgatgaccag agactacctg 1020  
 tcggtgaagg tgtgggacct caacatggag agcaggccgg tggagaccca ccaggtccac 1080  
 gagtacctgc gcagcaagct ctgctctctc tatgagaacg actgcatctt tgacaagttt 1140  
 gagtgttgct ggaacggttc ggatagcgcc atcatgaccg ggtcctataa caacttcttc 1200  
 aggatgtttg atagagacac gcggagggat gtgaccctgg aggcctcgag agagagcagc 1260  
 aaaccgcgcg ccagcctcaa accccggaag gtgtgtacgg ggggtaagcg gaggaagac 1320  
 gagatcagtg tggacagtct ggacttcaac aagaagatcc tgcacacagc ctggcacc 1380  
 gtggacaatg tcattgccgt ggctgccacc aataacttgt acatattcca ggacaaaatc 1440  
 aactagagac gcgaacgtga ggaccaagtc ttgtcttgca tagttaagcc ggacattttt 1500  
 ctgtcagaga aaaggcatca ttgtccgctc cattaagaac agtgacgcac ctgctacttc 1560  
 ccttcacaga cacaggagaa agccgcctcc gctggaggcc cgggtgtggt ccgcctcggc 1620  
 gaggcgcgag acaggcgctg ctgctcacgt ggagacgctc tcgaagcaga gttgacggac 1680  
 actgctccca aaaggtcatt actcagaata aatgtattta tttcaaaaaa aaaaaa 1737